TRANSACTIONS

OF THE

AMERICAN MATHEMATICAL SOCIETY

EDITED BY

ROBERT D. CARMICHAEL

FRANCIS R. SHARPE

JACOB D. TAMARKIN

VOLUME 37

JANUARY TO JUNE, 1935

PUBLISHED BY THE SOCIETY

MENASHA, WIS., AND NEW YORK

1935

General
Bound Aug, 1935

Composed, Printed and Bound by
The Collegiate Press
George Banta Publishing Company
Menasha, Wisconsin

TABLE OF CONTENTS

VOLUME 37, JANUARY TO JUNE, 1935

ON COMPLETE TOPOLOGICAL SPACES*

BY

JOHN von NEUMANN

Introduction

1. The notion of "completeness" is usually defined only for metric spaces (cf. for instance [1], p. 103). This seems reasonable, because this notion necessarily involves a certain "uniformity of the topology" of the space under consideration. Indeed, the definition of "completeness" is as follows:

Definition I. *If M is a space in which there is defined a metric* dist (f, g) *satisfying the usual postulates for distance* ([1], *p.* 94), *then a sequence F: $f_1, f_2, \cdots$ is fundamental if, for every $\delta > 0$, there exists an $n_1 = n_1(\delta)$ such that $m, n \geqq n_1$ imply* dist $(f_m, f_n) < \delta$; *and F is convergent if there exists an f such that, for every $\delta > 0$, there exists an $n_2 = n_2(\delta)$ such that $n \geqq n_2$ implies* dist $(f, f_n) < \delta$. *M is complete if every fundamental sequence is convergent.*

The need of uniformity in M arises from the fact that the elements of a fundamental sequence are postulated to be "near to each other," and not near to any fixed point. As a general topological space (cf. for instance [1], pp. 226–232) has no property which lends itself to the definition of such a "uniformity," it is improbable that a reasonable notion of "completeness" could be defined in it.

However, linear spaces (cf. [1], pp. 95–97, and Definition 1 in this paper), even if only topological, afford a possibility of "uniformization" for their topology: because of their homogeneity everything can be discussed in the neighborhood of 0. Thus one might introduce

Definition I′. *If L is a linear space (cf. above) with a topology (cf.* [1], *pp.* 226–232; *of course the "linear" operations αf [α a real number] and $f+g$ are supposed to be continuous), then a sequence $f_1, f_2, \cdots$ is fundamental if, for every neighborhood U of* 0 (*zero*), *there exists an $n_1 = n_1(U)$ such that $m, n \geqq n_1$ imply $f_m - f_n \epsilon U$;† and convergent if an f can be found so that for every neighborhood U of* 0 *there exists an $n_2 = n_2(U)$ such that $n \geqq n_2$ implies $f - f_n \epsilon U$. L is complete if every fundamental sequence is convergent.*

* Presented to the Society, December 28, 1934; received by the editors June 7, 1934.

† The fact that an element x belongs to a set S will be denoted by $x \epsilon S$ (not by $x \subset S$), while $T \subset S$ will mean that the set T is a subset of the set S. Other set-theoretical notations will be used: the sum of a set $(S, T, \cdots)$ of sets is $\mathfrak{S}(S, T, \cdots)$, the product (that is, the common part of the elements) of $(S, T, \cdots)$ is $\mathfrak{P}(S, T, \cdots)$, and the complementary set to S is $\mathfrak{C}S$. (These are not the notations of [1].)

This coincides with the previous definition if one defines, as usual, the neighborhoods of a point f_0 in the linear-metric case (in which dist $(f, g) \equiv$ dist $(f-g, 0) \equiv$ dist $(f-g)$) as spheres $S(f_0; \delta)$: dist $(f, f_0) =$ dist $(f-f_0) < \delta$, $\delta > 0$.

Another important notion is "total boundedness" (cf. [1], p. 108). We give the usual definition in the metric case and the generalization for the linear-topological case:

Definition II. *If M is a space in which a metric* dist (f, g) *is defined (cf. above), a set $S \subset M$ is totally bounded if, for every $\delta > 0$, there exist a finite number of spheres $S(f_1; \delta), \cdots, S(f_n; \delta)$ (of course $n, f_1, \cdots, f_n$ all depend on δ) such that $M \subset \mathfrak{S}(S(f_1; \delta), \cdots, S(f_n; \delta))$.*

Definition II′. *If L is a linear space with a topology (cf. above), a set $S \subset L$ is totally bounded if, for every neighborhood U of 0, a finite number of points $f_1, \cdots, f_n$ exist ($n, f_1, \cdots, f_n$ all depend on U) such that $M \subset \mathfrak{S}(f_1 + U, \cdots, f_n + U)$.**

Finally, we repeat the well known definition of "compactness" in a form which is particularly suited for our purposes.

Definition III. *If N is any topological space (cf. above) a set $S \subset N$ is compact if every infinite set $T \subset S$ has a condensation point*† *$f \epsilon S$. If we require only $f \epsilon N$, this expresses (at least if the countability axiom is satisfied) that S has a compact closure.*

An important fact connecting these notions is that I and II imply III (with $N = M$), the proof resulting from a simple application of the diagonal principle (cf. [1], pp. 108–109). The proof can be transferred immediately to the non-metric case: I′ and II′ imply III (with $N = L$), provided that the topology of L fulfills Hausdorff's first countability axiom (cf. [1], p. 229, axiom (9); for the proof, cf. Theorem 15 in this paper). That is: if L is complete and fulfills Hausdorff's first countability axiom, every totally bounded set $S \subset L$ has a compact closure.

2. It seems desirable, for various reasons, to get rid of the restriction represented by the countability axiom. Some important examples of linear spaces do not fulfill it.‡ Furthermore, the notions of total boundedness and closure-

* If L is a linear space, we use the following notation: ($f, g \epsilon L$; $S, T \subset L$; α, β real numbers): αS is the set of all $\alpha f, f \epsilon S$; $f \pm S$ is the set of all $f \pm g, g \epsilon S$; $S \pm T$ is the set of all $f \pm g, f \epsilon S$ and $g \epsilon T$. Note that $\alpha(S+T) = \alpha S + \alpha T$, $(\alpha\beta)S = \alpha(\beta S)$, and $S + T = T + S$, $(S+T)+R = S+(T+R)$; but only the weakened conditions $\alpha S + \beta S \supset (\alpha+\beta)S$, $(S \pm T) \mp T \supset S$ are valid.

† f is a condensation point of T if $\mathfrak{P}(T, U)$ is infinite for every neighborhood U of f.

‡ For example, Hilbert space in its "weak" topology (cf. for instance [2], p. 379); the space of all bounded operators in Hilbert space, in its "strong" and in its "weak" topology (cf. [2], pp. 381–382; for the discussion of all these topologies, [2], pp. 378–388).

compactness play an important role in the general theory of almost periodic functions (cf. the following paper of S. Bochner and J. von Neumann on this subject), and their equivalence is necessary for the smooth working of this theory, which makes no other use of the countability axiom, and which therefore should be workable without its help. (This has actually been done, loc. cit., by the use of the results of this paper.) But if we use the definition of completeness given in I′, then II′ does not necessarily imply III (that is, total boundedness does not imply closure-compactness) if the countability axiom does not hold. Therefore we have to find another definition of completeness which leads to the desired implication.

The simplest thing is to postulate this directly:

DEFINITION IV. *If L is a linear space with a topology, it is topologically complete if every totally bounded set $S \subset L$ has a compact closure.*

For metric linear spaces, and even for every linear space satisfying the countability axiom, this is equivalent to the usual definition of completeness (I or I′; cf. Theorem 15). The various spaces mentioned at the beginning of this chapter are topologically complete (cf. Theorem 23). The most important property of this notion is, however, that if L is topologically complete, the linear space formed by the functions with a given domain D and with a range $\subset L$ is (if subjected to certain restrictions, like boundedness, etc., cf. Definition 11) topologically complete too. This is rather obvious for the Definitions I and I′, but not at all for IV; we will prove it in Theorem 18. All these properties make our notion of topological completeness just as useful for various applications (for instance in the generalized theory of almost periodic functions, as mentioned above), as the usual notion of (metric) completeness, while its range of generality is essentially wider.

We now pass on to the exact exposition of the subject.

I. DEFINITIONS

3. We define linear spaces in the usual way (cf. for instance [1], pp. 95–97):

DEFINITION 1. *The set L is a linear space, if, for $f, g \in L$ and any real number α,* αf and $f+g$ are in L and are defined so that*

(1) $f + g = g + f,$ (2) $(f + g) + h = f + (g + h),$

(3) $1 \cdot f = f,$ (4) $\alpha(\beta f) = (\alpha\beta)f,$

(5) $(\alpha + \beta)f = \alpha f + \beta f,$ (6) $\alpha(f + g) = \alpha f + \alpha g,$

(7) $f + h = g + h$ *implies* $f = g$.

* It would be sufficient to admit only rational α's.

By (3) and (5), $f+0\cdot f=f$; by (1) and (2), $(f+g)+0\cdot f=f+g$; by interchanging f and g, $(f+g)+0\cdot g=f+g$; by (7), $0\cdot f=0\cdot g$. Thus $0\cdot f$ is independent of f; and we call it 0 (zero). We have $f+0=f$. Writing $-f$ for $(-1)\cdot f$, (5) gives $f+(-f)=0$; writing $f-g$ for $f+(-g)$, (1) and (2) give $(f-g)+g=f$; and by (7), $x=f-g$ is the only solution of $x+g=f$. Now all rules of computation for 0, $-f$, $f-g$ are easily deduced.

A metric (or an "absolute value") in L is defined in the usual way (cf. [1], p. 97):

DEFINITION 2a. *The linear set L is metric if, for every $f\epsilon L$, a real number $\|f\|$, its "absolute value," is defined, such that*

(1) $\|f\| > 0$ *if* $f \neq 0$, (2) $\|\alpha f\| = |\alpha| \cdot \|f\|$, (3) $\|f+g\| \leqq \|f\| + \|g\|$.

The metric is then defined by dist $(f, g)=\|f-g\|$.

This dist (f, g) possesses the characteristic properties of a distance and can be used to define a topology in L (cf. [1], p. 94; also the end of paragraph 1 of this paper). However, we shall not assume that L is metric, but only that it has a topology. This is done in the following definition, in which it was attempted to reduce the strength of the postulates to the necessary minimum.

DEFINITION 2b. *The linear set L is topological if a set $\mathfrak{U}$ of sets $U \subset L$ is given such that*

(1) *if $U \epsilon \mathfrak{U}$, then $0 \epsilon U$,*
(2) *there is a sequence $U_1, U_2, \cdots \epsilon \mathfrak{U}$ such that $\mathfrak{P}(U_1, U_2, \cdots)=(0)$,*†‡
(3) *if $U, V \epsilon \mathfrak{U}$, there is a $W \epsilon \mathfrak{U}$ with $W \subset \mathfrak{P}(U, V)$,*‡
(4) *if $U \epsilon \mathfrak{U}$, there is a $V \epsilon \mathfrak{U}$ such that, for every α with $-1 \leqq \alpha \leqq 1$, $\alpha V \subset U$,*§
(5) *if $U \epsilon \mathfrak{U}$, there is a $V \epsilon \mathfrak{U}$ with $V+V \subset U$,*§
(6) *if $f \epsilon L$, $U \epsilon \mathfrak{U}$, there is an α with $f \epsilon \alpha U$.*§

L is "convex" if the further condition

(7) *if $U \epsilon \mathfrak{U}$, then $U+U \subset 2U$*§

is fulfilled.

† If Hausdorff's first countability axiom holds, (2) is fulfilled by choosing a complete system of neighborhoods of 0 for $U_1, U_2, \cdots$ (cf. [1], p. 229, axiom (9)), but the converse is not true: (2) is essentially weaker than the countability axiom. This is shown by the examples of Part IV, Theorem 23; cf. [3], p. 264.

‡ (2) and (3) could be replaced by two other postulates (2′) and (3′) which extend (2) and restrict (3):

(2′) there is an aleph $\aleph^*$ and a set $(U, V, \cdots) \subset \mathfrak{U}$ with this aleph $\aleph^*$, such that $\mathfrak{P}(U, V, \cdots)=(0)$;

(3′) for each set $(U, V, \cdots) \subset \mathfrak{U}$ with an aleph $< \aleph^*$ there is a $W \epsilon \mathfrak{U}$ with $W \subset \mathfrak{P}(U, V, \cdots)$.

All our discussions could be carried through, with little change, on this basis. In the present form of (2) and (3), $\aleph^* = \aleph_0$ (the set $(U, V, \cdots)$ is countable).

§ See first footnote on p. 2.

If L is metric, we consider the spheres ($\delta > 0$)

$$S^{\circ}(f_0;\delta)\text{: the set of all } f \text{ with } \|f - f_0\| < \delta,$$
$$S^{\circ l}(f_0;\delta)\text{: the set of all } f \text{ with } \|f - f_0\| \leqq \delta.$$

Choosing $\mathfrak{U}$ as the set of all $S^{\circ}(0;\delta)$ or as the set of all $S^{\circ l}(0;\delta)$ makes L topological and convex (one easily verifies Definition 2b, (1)–(7)), and the topology, which we will define with the aid of $\mathfrak{U}$ in Definition 4 and Theorem 6, coincides in this case with the usual metric topology of L.

II. General theorems

4. In the following discussion of Chapters II and III, L is assumed merely to be a topological space, that is, to fulfill Definition 2b, (1)–(6), except where the contrary is expressly stated.

THEOREM 1. *If* $U \epsilon \mathfrak{U}$, $A > 0$, $n = 1, 2, \cdots$, *there is, for each value of* n, *a* $V \epsilon \mathfrak{U}$ *such that, for all sets* $\alpha_1, \cdots, \alpha_n$ *with* $-A \leqq \alpha_1, \cdots, \alpha_n \leqq A$, $\alpha_1 V + \cdots + \alpha_n V \subset U$.

Increasing A and n strengthens the statement, so we may assume $A = 2^p$, $n = 2^q$, $p, q = 0, 1, 2, \cdots$. It is sufficient to obtain $AV + \cdots + AV \subset U$ (n addends), because the W of Definition 2b, 4 (with $\alpha W \subset V$ for $-1 \leqq \alpha \leqq 1$) will then have the desired properties. As $AV \subset V + \cdots + V$ (A addends),* the statement is strengthened if we replace A, n by 1, $An = 2^{p+q}$, respectively. Thus we may assume $A = 1$, $n = 2^r$, $r = 0, 1, 2, \cdots$. We have to find a $V \epsilon \mathfrak{U}$ with $V + \cdots + V \subset U$ (2^r addends).

For $r = 0$ we may choose $V = U$; if we have a V for any $r = 0, 1, 2, \cdots$, we can apply Definition 2b, (5), to it, and thus obtain a V for $r+1$. This completes the proof.

DEFINITION 3. *If* $S \subset L$, S_i *is the set of all* f *for which a* $U \epsilon \mathfrak{U}$ *with* $f + U \subset S$ *exists.*†

THEOREM 2. $S_{ii} = S_i \subset S$.

$0 \epsilon U$, $f \epsilon f + U$, so that $S_i \subset S$. Therefore $S_{ii} \subset S_i$. Now if $f \epsilon S_i$, that is, if $f + U \subset S$, choose a $V \epsilon \mathfrak{U}$ with $V + V \subset U$ (Definition 2b, (5)). Then for $g \epsilon f + V$, $g + V \subset f + V + V \subset f + U \subset S$, and $g \epsilon S_i$. Thus $f + V \subset S_i$, $f \epsilon S_{ii}$, and therefore $S_i \subset S_{ii}$. This completes the proof.

THEOREM 3. $0 \epsilon U_i$; $(f+S)_i = f + S_i$; *if* $\alpha \neq 0$, $(\alpha S)_i = \alpha S_i$; $S_i + T_i \subset (S+T)_i$.

The first two statements are obvious. If $f \epsilon S_i$ and $f + U \subset S$, then $\alpha f + \alpha U \subset \alpha S$, and if V is chosen by Theorem 1 with $V/\alpha \subset U$, $\alpha f + V \subset \alpha S$ and

* See first footnote on p. 2.

† S_i is the set of inner points of S (cf. Theorem 4).

$\alpha f \epsilon (\alpha S)_i$. Thus $\alpha S_i \subset (\alpha S)_i$; substitution of $1/\alpha$, αS for α, S leads to the result that $(\alpha S_i) \subset \alpha S_i$, proving the third statement. If $f \epsilon S_i$ and $g \epsilon T_i$, $f+g \epsilon f+T_i = (f+T)_i \subset (S+T)_i$, proving the fourth statement.

Definition 4. *S is open if $S=S_i$; S is closed if $\mathfrak{C}S$* is open. This means that if S is closed, $S=S_{cl}$, where we define $S_{cl}=\mathfrak{C}((\mathfrak{C}S)_i)$.*†

Theorem 4. *S_i is open and the greatest open set $\subset S$, S_{cl} is closed and the smallest closed set $\supset S$.*

The statements about S_i follow from Theorems 2 and 3; those about S_{cl} result from considering $\mathfrak{C}S$.

Theorem 5. *For every S, $S_{cl} = \mathfrak{P}(S+U)$, where U runs over all elements of* $\mathfrak{U}$.

If $U \epsilon \mathfrak{U}$, there is a $V \epsilon \mathfrak{U}$ with $-V \subset U$ (Definition 2b, (4)), and thus $V \subset -U$. For this reason $\mathfrak{P}(S+U) = \mathfrak{P}(S-V)$ (U, V run over all $\mathfrak{U}$). Now $f \epsilon \mathfrak{P}(S-V)$ means that for every $V \epsilon \mathfrak{U}$, $f \epsilon g-V$ for some $g \epsilon S$, that is, $g \epsilon S$, $g \epsilon f+V$. But this is equivalent to $f \epsilon S_{cl}$.

Theorem 6. *In the sense in which Hausdorff defined a topology, using the open sets as fundamental notions (cf.* [1], *p.* 228), *Definition* 4 *describes a regular Hausdorff topology, that is, one which fulfills Hausdorff's axioms* (1)–(3) *and* (6) (*cf.* [1], *pp.* 228–229; *thus all axioms* (1)–(6) *are fulfilled*).

Ad (1): 0 and L are obviously open. Ad (2): If S_1, S_2 are open, assume $f \epsilon \mathfrak{P}(S_1, S_2)$. Then $f+U_1 \subset S_1$, $f+U_2 \subset S_2$, and so with $V \epsilon \mathfrak{U}$, $V \subset \mathfrak{P}(U_1, U_2)$ (by Definition 2b, (3)) and $f+V \subset \mathfrak{P}(S_1, S_2)$, proving that $f \epsilon \mathfrak{P}(S_1, S_2)_i$. Thus $\mathfrak{P}(S_1, S_2)$ is open. Ad (3): If $S, T, \cdots$ are open, $\mathfrak{S}(S, T, \cdots)$ is obviously open. Ad (6): Let $\bar{S}$ be closed, f not an element of $\bar{S}$. Then $\mathfrak{C}\bar{S}$ is open, $f \epsilon \mathfrak{C}\bar{S}$, so that $U \epsilon \mathfrak{U}$, $f+U \subset \mathfrak{C}\bar{S}$. Choosing $V \epsilon \mathfrak{U}$ with $V+V \subset U$ (by Definition 2b, (3)) we have $(f+V_i)_{cl} \subset (f+V)_{cl} \subset f+V+V$ (by Theorem 5) $\subset f+U \subset \mathfrak{C}\bar{S}$, that is, for $T=f+V_i$, T is open, $f \epsilon T$, and $\mathfrak{P}(T_{cl}, \bar{S})$ is empty.

In the following discussions we shall always consider L as topologized by the topology of Definition 4 and Theorem 6, except where the contrary is explicitly stated. Thus we can use the whole topological terminology: we can speak of open and closed sets, which have already been defined in harmony with this by Definition 4, of continuous functions, limits of sequences, condensation points, etc.

Theorem 7. *αf and $f+g$ are continuous functions of α, f and f, g respectively.*

---

* See second footnote on p. 1.

† S_{cl}, the closure of S, is the set of all points and all condensation points of S (cf. Theorem 4).

Ad $f+g$: Assume $f_0+g_0 \epsilon S$, S open. Choose $U \epsilon \mathfrak{U}$, $f_0+g_0+U \subset S$, $V \epsilon \mathfrak{U}$, $V+V \subset U$. Then

$$(f_0 + V_i) + (g_0 + V_i) \subset f_0 + g_0 + V_i + V_i \subset f_0 + g_0 + V + V \subset f_0 + g_0 + U \subset S,$$

that is, the sets $T_1 = f_0 + V_i$, $T_2 = g_0 + V_i$ are open, $f_0 \epsilon T_1$, $g_0 \epsilon T_2$, $T_1 + T_2 \subset S$. Ad αf: Due to the continuity of $f+g$ (and Definition 1, (5)–(6)) we need consider only $\alpha_0 = 0$ or $f_0 = 0$. Then, in any event, $\alpha_0 f_0 = 0$, so that we have the situation $0 \epsilon S$, S open. Ad $\alpha_0 = 0$: Choose $U \epsilon \mathfrak{U}$, $U \subset S$, and $V \epsilon \mathfrak{U}$ by Theorem 1 with $n=2$, $A=1$. Now choose a β with $f_0 \epsilon \beta V$ (by Definition 2b, (6)); then for

$$|\alpha| < \frac{1}{\max(1, |\beta|)}$$

we have $\alpha(f_0+V) \subset \alpha\beta V + \alpha V \subset U \subset S$, that is, the sets

$$A: \qquad |\alpha| < \frac{1}{\max(1, |\beta|)}$$

and $T = f_0 + V_i$ are open, $0 \epsilon A$, $f_0 \epsilon T$, $AT \subset S$. Ad $f_0 = 0$: Choose $U \epsilon \mathfrak{U}$, $U \subset S$ and $V \epsilon \mathfrak{U}$ by Theorem 1 with $n=1$, $A = |\alpha_0| + 1$. Then $|\alpha - \alpha_0| < 1$ implies $|\alpha| \leqq A$, $\alpha V \subset U \subset S$, that is, the sets A: $|\alpha - \alpha_0| < 1$ and $T = V_i$ are open, $\alpha_0 \epsilon A$, $0 \epsilon T$, $AT \subset S$.

5. We now introduce the notion of boundedness, which is usually considered as a metric notion. One sees at once, remembering the remarks made at the end of §3, that our general definition coincides in the case of a metric L with the usual definition.

Definition 5. *S is bounded if, for every $U \epsilon \mathfrak{U}$, there is an α such that $S \subset \alpha U$.*

Remark. We could replace herein the $U \epsilon \mathfrak{U}$ by the open sets T with $0 \epsilon T$ for if T is such a set, a $U \epsilon \mathfrak{U}$, $U \subset T$, exists, and for $U \epsilon \mathfrak{U}$ an open T, $0 \epsilon T$, $T \subset U$ exists: $T = U_i$.

Theorem 8. *Every finite set is bounded. If $S, \cdots, T$ are a finite number of bounded sets, $\mathfrak{S}(S, \cdots, T)$ is bounded. If S, T are bounded sets, αS, $f+S$, $S+T$ are bounded.*

It follows by Definition 2b, (6), that a one-element set (f) is bounded; therefore every finite set is bounded if the second statement is true. If the second statement holds for two addends, it holds by induction for any finite number. Let S, T be bounded, $U \epsilon \mathfrak{U}$, choose $V \epsilon \mathfrak{U}$ by Definition 2b, (4), and α, β with $S \subset \alpha V$, $T \subset \beta V$. Then for $\gamma = \max(|\alpha|, |\beta|)$,

$$S \subset \alpha V = \gamma\left(\frac{\alpha}{\gamma} V\right) \subset \gamma U;$$

similarly, $T \subset \gamma U$, $\mathfrak{S}(S, T) \subset \gamma U$. Thus the first two statements are proved.

In the last statement the first part (concerning αS) is obvious; the second follows from the third by putting $T=(f)$, so that we have to prove only the third part. Assume $U \in \mathfrak{U}$, choose $V \in \mathfrak{U}$ by Theorem 1 with $n=2$, $A=1$. Choose β, γ with $S \subset \beta V$, $T \subset \gamma V$. Then, for $\delta = \max(|\beta|, |\gamma|)$,

$$S + T \subset \beta V + \gamma V = \delta\left(\frac{\beta}{\delta} V\right) + \delta\left(\frac{\gamma}{\delta} V\right) \subset \delta U.$$

This completes the proof.

The next definition is a repetition of Definition II′ in §1 (cf. the comment given there).

DEFINITION 6. *S is totally bounded if, for every $U \in \mathfrak{U}$, there is a finite number of elements $f_1, \cdots, f_n$ of L, such that $S \subset \mathfrak{S}(f_1+U, \cdots, f_n+U)$.*

Remark 1. The $f_1, \cdots, f_n$ could be restricted to S. In this form the condition is obviously sufficient; but it is also necessary: Choose $V \in \mathfrak{U}$ by Theorem 1, with $n=2$, $A=1$; then $V-V \subset U$. Apply Definition 6 to V: $S \subset \mathfrak{S}(g_1+V, \cdots, g_n+V)$. As the sets $g_\nu+V$ which contain no point of S can be omitted, we may assume that every $g_\nu+V$ contains a point of S, say f_ν. Then $g_\nu \in f_\nu - V$, $S \subset \mathfrak{S}(f_1-V+V, \cdots, f_n-V+V) \subset \mathfrak{S}(f_1+U, \cdots, f_n+U)$.

Remark 2. For the reasons given in the Remark after Definition 5, we could replace the $U \in \mathfrak{U}$ by the open sets T with $0 \in T$ in all these considerations.

Remark 3. The set of all real numbers is a particularly simple linear space. It is clear that its customary metric, $\|x\|$ = absolute value of x, is a metric in the sense of Definition 2a, and a topology in the sense of Definition 2b (cf. the end of Part I). The boundedness in the sense of Definition 5, and the total boundedness in the sense of Definition 6, obviously coincide with the customary notion of boundedness for real numbers in this case.

THEOREM 9. *Every finite set is totally bounded. The set of all αf, $-1 \leqq \alpha \leqq 1$ (f fixed), is totally bounded. If S, T are totally bounded, αS, $f+S$, $S+T$ are also totally bounded, and so is $\mathfrak{S}(S, T)$.*

If $L_1, \cdots, L_k$, L are topological linear spaces, if $S_\kappa \subset L_\kappa$ and is totally bounded, $\kappa=1, \cdots, k$, if $\mathfrak{F}(f_1, \cdots, f_k)$ is a function with the domain $f_\kappa \in S_\kappa$, $\kappa=1, \cdots, k$, uniformly continuous in this domain, and with a range $\subset L$, then the set of all $\mathfrak{F}(f_1, \cdots, f_k)$, $f_\kappa \in S_\kappa$, $\kappa=1, \cdots, k$, is totally bounded.

The statements for finite sets, $f+S$, $\mathfrak{S}(S, T)$, are obvious, and the statement for αS follows from Theorem 1. The statements concerning the sets αf, $-1 \leqq \alpha \leqq 1$, and $S+T$ are both special cases of the last statement ($k=1$, L = set of all real numbers, $L_1=L$, f_1 is to be replaced by α, $\mathfrak{F}(\alpha)=\alpha f$; and $k=2$, $L_1=L_2=L$, $\mathfrak{F}(f_1, f_2)=f_1+f_2$ respectively; cf. Theorem 7 and Remark

3 after Definition 6). So our only task is to prove that statement. Denote the $\mathfrak{U}$ of $L_1, \cdots, L_k, L$ by $\mathfrak{U}_1, \cdots, \mathfrak{U}_k, \mathfrak{U}$ respectively. If a $U \epsilon \mathfrak{U}$ is given, the uniform continuity of $\mathfrak{F}$ means that we can choose $U_1 \epsilon \mathfrak{U}_1, \cdots, U_k \epsilon \mathfrak{U}_k$ so that the conditions $f_\kappa - g_\kappa \epsilon U_\kappa$, f_κ, $g_\kappa \epsilon S_\kappa$, $\kappa = 1, \cdots, k$, imply that $\mathfrak{F}(f_1, \cdots, f_k) - \mathfrak{F}(g_1, \cdots, g_k) \epsilon U$. Now apply Definition 6 and its Remark 1 to S_κ: $S_\kappa \subset \mathfrak{S}(g_{\kappa,1} + U_\kappa, \cdots, g_{\kappa,n_\kappa} + U_\kappa)$, $g_{\kappa,1}, \cdots, g_{\kappa,n_\kappa} \epsilon S_\kappa$. If a system $f_\kappa \epsilon S_\kappa$, $\kappa = 1, \cdots, k$, is given, we have $f_\kappa \epsilon g_{\kappa,\nu_\kappa} + U_\kappa$, with a $\nu_\kappa = 1, \cdots, n_\kappa$ for each $\kappa = 1, \cdots, k$, and thus $\mathfrak{F}(f_1, \cdots, f_k) \epsilon \mathfrak{F}(g_{1,\nu_1}, \cdots, g_{k,\nu_k}) + U$. So if we put $n = n_1 \cdots n_k$, and arrange the $\mathfrak{F}(g_{1,\nu_1}, \cdots, g_{k,\nu_k})$ ($\nu_\kappa = 1, \cdots, n_\kappa$, $\kappa = 1, \cdots, k$) in some order $\mathfrak{F}_1, \cdots, \mathfrak{F}_n$, then our set is contained in $\mathfrak{S}(\mathfrak{F}_1 + U, \cdots, \mathfrak{F}_n + U)$.

THEOREM 10. *Every totally bounded set S is bounded.*

Let S be totally bounded, $U \epsilon \mathfrak{U}$. Choose $V \epsilon \mathfrak{U}$ by Theorem 1, with $n = 2$, $A = 1$, and then $f_1, \cdots, f_n$ with $S \subset \mathfrak{S}(f_1 + V, \cdots, f_n + V)$. Select $\alpha_1, \cdots, \alpha_n$ with $f_\nu \epsilon \alpha_\nu V$ and put $\beta = \max(1, |\alpha_1|, \cdots, |\alpha_n|)$; then

$$f_\nu + V \subset \alpha_\nu V + V = \beta\left(\frac{\alpha_\nu}{\beta} V\right) + \beta\left(\frac{1}{\beta} V\right) \subset \beta U.$$

THEOREM 11. *The boundedness (total boundedness) of S is a necessary and sufficient condition for that of S_{cl}.*

As $S \subset S_{\mathrm{cl}}$, the condition is necessary. If $U \epsilon \mathfrak{U}$, choose $V \epsilon \mathfrak{U}$, $V + V \epsilon \mathfrak{U}$; then by Theorem 5, $V_{\mathrm{cl}} \subset V + V \subset U$, and thus $S \subset \alpha V$ (or $S \subset \mathfrak{S}(f_1 + V, \cdots, f_n + V)$) implies $S_{\mathrm{cl}} \subset \alpha U$ (or $S_{\mathrm{cl}} \subset \mathfrak{S}(f_1 + U, \cdots, f_n + U)$). Thus the condition is also sufficient.

We now investigate convexity.

THEOREM 12. *If L is convex (Definition 2b, (7)) then, for $U \epsilon \mathfrak{U}$ and $\alpha_1, \cdots, \alpha_n \geqq 0$, $\alpha_1 U_{\mathrm{cl}} + \cdots + \alpha_n U_{\mathrm{cl}} = (\alpha_1 + \cdots + \alpha_n) U_{\mathrm{cl}}$.*

Induction proves this theorem for all $n = 1, 2, \cdots$, if it holds for $n = 2$. For $\alpha_1 = \alpha_2 = 0$ it is obvious, therefore we may assume $\alpha_1 + \alpha_2 > 0$. Division by $\alpha_1 + \alpha_2$ then gives

$$\alpha U_{\mathrm{cl}} + (1 - \alpha) U_{\mathrm{cl}} = U_{\mathrm{cl}} \qquad \left(\alpha = \frac{\alpha_1}{\alpha_1 + \alpha_2},\ 0 \leqq \alpha \leqq 1\right).$$

As $\supset$ is obvious, we need to prove only $\subset$.

Now Definition 2b, (7), states that $U + U \subset 2U$; iterated n times, this becomes $U + \cdots + U$ (2^n addends) $\subset 2^n U$, and, a fortiori, $kU + (2^n - k)U \subset 2^n U$, $k = 0, 1, \cdots, 2^n$. Thus $\alpha U + (1 - \alpha) U \subset U$ if $0 \leqq \alpha \leqq 1$ with α dyadic-rational; from this, consideration of continuity leads to $\alpha U_{\mathrm{cl}} + (1 - \alpha) U_{\mathrm{cl}} \subset U_{\mathrm{cl}}$ for all $0 \leqq \alpha \leqq 1$, completing the proof.

DEFINITION 7. *S_{conv} is the set of all $\alpha_1 f_1 + \cdots + \alpha_n f_n$, with $n = 1, 2, \cdots$, $\alpha_1, \cdots, \alpha_n \geqq 0$, $\alpha_1 + \cdots + \alpha_n = 1$, $f_1, \cdots, f_n \epsilon S$.*

THEOREM 13. *If L is convex, then, for all $U \epsilon \mathfrak{U}$, $(U_{cl})_{conv} = U_{cl}$.*

This follows immediately from Theorem 12.

THEOREM 14. *The boundedness (total boundedness) of S is a necessary condition for that of S_{conv}; if L is convex, it is also sufficient.*

As $S \subset S_{conv}$, the condition is necessary. For the sufficiency we have to assume that L is convex. If $U \epsilon \mathfrak{U}$, choose $V \epsilon \mathfrak{U}$, $V + V \subset U$; then $V_{conv} \subset (V_{cl})_{conv} = V_{cl} \subset V + V \subset U$ (by Theorems 5, 13). Thus $S \subset \alpha V$ implies $S_{conv} \subset \alpha V_{conv} \subset \alpha U$, settling the case for boundedness. $S \subset \mathfrak{S}(f_1 + V, \cdots, f_n + V)$ implies $S_{conv} \subset \mathfrak{S}(\beta_1 f_1 + \cdots + \beta_n f_n + V_{conv}) \subset \mathfrak{S}(\beta_1 f_1 + \cdots + \beta_n f_n + U)$, where the $\beta_1, \cdots, \beta_n$ run over all combinations with $\beta_1, \cdots, \beta_n \geqq 0$, $\beta_1 + \cdots + \beta_n = 1$. If the set of these $\beta_1 f_1 + \cdots + \beta_n f_n$ is totally bounded, then this set is contained in $\mathfrak{S}(g_1 + U, \cdots, g_p + U)$, and thus $S_{conv} \subset \mathfrak{S}(g_1 + U + U, \cdots, g_p + U + U)$. We could replace U by V, and we would then have $S_{conv} \subset \mathfrak{S}(g_1 + U, \cdots, g_p + U)$. This settles the case for total boundedness too, provided that the set of the $\beta_1 f_1 + \cdots + \beta_n f_n$ (n and $f_1, \cdots, f_n$ fixed) is totally bounded.

This is a subset of the $(\beta_1 f_1 + \cdots + \beta_n f_n)$-set with $0 \leqq \beta_1 \leqq 1, \cdots, 0 \leqq \beta_n \leqq 1$. This set is disposed of by Theorem 9, if each set $\beta_\nu f_\nu$, $0 \leqq \beta_\nu \leqq 1$, is totally bounded, but this again follows from Theorem 9.

Finally we repeat Definition III in §1.

DEFINITION 8. *S is compact if every infinite set $T \subset S$ has a condensation point $f \epsilon S$.*

Note that we did not assume that any of Hausdorff's countability axioms hold (cf. [1], p. 229, axioms (9) and (10)), and that we still are considering "compactness" and not the Alexandroff-Urysohn "bicompactness" (cf. [3], [4], in particular [3], pp. 259–260), although the latter is specially adapted to these cases. The reason is that for the totally bounded sets S compactness implies the countability axioms (although they need not hold for L, cf. Theorem 16) and thus bicompactness.

III. TOPOLOGICAL COMPLETENESS

6. The two definitions of completeness which we discussed in §1, I', and IV, are the following:

DEFINITION 9. *L is sequentially complete if every "fundamental sequence" $f_1, f_2, \cdots$ in L (i.e., every sequence such that for each $U \epsilon \mathfrak{U}$ there is an $n_1 = n_1(U)$, such that $m, n \geqq n_1$ imply $f_m - f_n \epsilon U$) is "convergent" (i.e., an f exists such that for each $U \epsilon \mathfrak{U}$ there is an $n_2 = n_2(U)$, such that $n \geqq n_2$ implies $f_n - f \epsilon U$).*

DEFINITION 10. *L is topologically complete if every closed and totally bounded set $S \subset L$ is compact.*

To characterize the relationship between these two notions, we prove

THEOREM 15. *Sequential completeness is a necessary condition for topological completeness; if L satisfies Hausdorff's first countability axiom (cf. above), it is also sufficient.*

Assume first that L is topologically complete, and let $f_1, f_2, \cdots$ be a fundamental sequence. If $U \epsilon \mathfrak{U}$, $f_m \epsilon f_{N_1} + U$ for $m \geqq N_1 = N_1(U)$; thus $(f_1, f_2, \cdots) \subset \mathfrak{S}(f_1 + U, \cdots, f_{N_1} + U)$. So $(f_1, f_2, \cdots)$ is totally bounded; $(f_1, f_2, \cdots)_{cl}$ is also totally bounded (by Theorem 11) and closed, and thus compact. If infinitely many $f_1, f_2, \cdots$ are different, $(f_1, f_2, \cdots)$ is an infinite set $\subset (f_1, f_2, \cdots)_{cl}$, and has a condensation point f, that is, an f such that, for each $U \epsilon \mathfrak{U}$, there are infinitely many n for which $f_n \epsilon f + U$. If only a finite number of $f_1, f_2, \cdots$ are different, infinitely many f_n must coincide, and their common value f has the above property. So such an f exists in any event. Corresponding to $U \epsilon \mathfrak{U}$, choose $V \epsilon \mathfrak{U}$, $V + V \subset U$, $n_1 = n_1(V)$, and the above f with respect to V. Then $f_n - f \epsilon V$ occurs for some $n \geqq n_1$, and $f_m - f_n \epsilon V$ for every $m, n \geqq n_1$; thus, for every $m \geqq n_1$, $f_m - f = (f_m - f_n) + (f_n - f) \epsilon V + V \subset U$. Putting $n_2(U) = n_1(V)$ we see that $f_1, f_2, \cdots$ is convergent, and thus L is sequentially complete.

Now consider the converse situation. Let L be sequentially complete, let $U_1, U_2, \cdots$ be a complete system of neighborhoods for 0 (cf. [1], p. 229; this means that for each $U \epsilon \mathfrak{U}$ some $U_n \subset U$), and assume S to be totally bounded and closed. As every infinite $T \subset S$ contains a sequence $f_1, f_2, \cdots$ of distinct elements, we may assume $T = (f_1, f_2, \cdots)$. As every fundamental sequence is convergent and thus has a condensation point f, we need only exhibit a fundamental subsequence $f^{(1)}, f^{(2)}, \cdots$ in a sequence $(f_1, f_2, \cdots) \subset S$. Now we can apply the well known diagonal process.

Choose $g_{n1}, \cdots, g_{nm_n}$ with $S \subset (g_{n1} + U_n, \cdots, g_{nm_n} + U_n)$. Some $g_{1\nu} + U_1$, say $\nu = \nu_1$, contains infinitely many elements $f_1^{(1)}, f_2^{(1)}, \cdots$ of the sequence $f_1, f_2, \cdots$. Some $g_{2\nu} + U_2$, say $\nu = \nu_2$, contains infinitely many elements $f_1^{(2)}, f_2^{(2)}, \cdots$ of the sequence $f_1^{(1)}, f_2^{(1)}, \cdots$. And so on. Now put $f^{(1)} = f_1^{(1)}$, $f^{(2)} = f_2^{(2)}, \cdots$. If $U \epsilon \mathfrak{U}$, choose $V \epsilon \mathfrak{U}$, $V - V \subset U$ (cf. Definition 2b, (4), (5)) *and* $U_p \subset V$. For $m \geqq p$, $f^{(m)} = f_m^{(m)}$ belongs to $(f_1^{(m)}, f_2^{(m)}, \cdots)$, thus to $(f_1^{(p)}, f_2^{(p)}, \cdots)$, and to $g_{p\nu_p} + U_p$. Thus, for $m, n \geqq p$, $f^{(m)} - f^{(n)} \epsilon U_p - U_p \subset V - V \subset U$. Putting $n_1(U) = p$ we see that $f^{(1)}, f^{(2)}, \cdots$ is fundamental, thus completing the proof of the topological completeness of L.

As our main interest belongs to the cases in which L violates the countability axiom, and as the equivalence of sequential and of topological com-

pleteness has not been established for them, we continue by investigating the properties of topological completeness.

THEOREM 16. *If L is topologically complete and S totally bounded, then if we consider S as a space with the topology of L, this is a normal and separable Hausdorff topology in S, that is, one which fulfills Hausdorff's axioms* (1)–(3), (8), *and* (10) (*cf.* [1], *pp.* 228–229; *thus all axioms* (1)–(10) *are fulfilled*).

Remark. For this reason we can replace condensation points in S by limits of convergent sequences.—

As $S \subset S_{cl}$ and as S_{cl} is also totally bounded, we may consider S_{cl} instead of S, that is, we can restrict ourselves to closed sets S. Then S is compact. The topology of L fulfilled axioms (1)–(3), (6) in L, therefore it also fulfills them in S. Let us therefore consider the other axioms.

Ad (9): Choose $U_1, U_2, \cdots$ by Definition 2b, (2), with $\mathfrak{P}(U_1, U_2, \cdots) = (0)$. Now choose (by Definition 2b, (3), (5)) $V_n \epsilon \mathfrak{U}$, $V_n + V_n \subset U_n$ and $W_n \epsilon \mathfrak{U}$, $W_n \subset \mathfrak{P}(V_1, \cdots, V_n)$. Let T be an open set, $0 \epsilon T$. Assume that, for $n = 1, 2, \cdots, \mathfrak{P}(W_n, S)$ is not a subset of $\mathfrak{P}(T, S)$. Then there exists an f_n with $f_n \epsilon \mathfrak{P}(W_n, S)$, f_n not an element of $\mathfrak{P}(T, S)$. If infinitely many $f_1, f_2, \cdots$ are different, $(f_1, f_2, \cdots)$ is an infinite set $\subset S$, and has a condensation point f, that is, an f such that, for each $U \epsilon \mathfrak{U}$, there are infinitely many m for which $f_m \epsilon f + U$.If only a finite number of $f_1, f_2, \cdots$ are different, infinitely many f_n must coincide, and their common value f has the above property. So such an f exists in any event. Choose $m \geqq n$ and $f_m \epsilon f + U$. $f_m \epsilon W_m \subset V_n$, $f_m \epsilon \mathfrak{C}T$, so that f is a point or a condensation point of V_n and of $\mathfrak{C}T$; hence $f \epsilon (V_n)_{cl} \subset V_n + V_n \subset U_n$ (by Theorem 5), and $f \epsilon \mathfrak{C}T$ ($\mathfrak{C}T$ is closed). Thus $f \epsilon \mathfrak{P}(U_1, U_2, \cdots) = (0)$, $f = 0 \epsilon T$, contradicting the condition that $f \epsilon \mathfrak{C}T$. This proves that an $n = 1, 2, \cdots$ with $\mathfrak{P}(W_n, S) \subset \mathfrak{P}(T, S)$ must exist, so that $(W_1)_i, (W_2)_i, \cdots$ form a complete system of neighborhoods of 0 in S, and $f + (W_1)_i, f + (W_2)_i, \cdots$ form a complete system of neighborhoods of f in S (consider 0 in $-f + S$).

Ad (10): Take the $W_1, W_2, \cdots$ constructed above and choose $X_n \epsilon \mathfrak{U}$, $X_n - X_n \subset W_n$ (by Theorem 1, $n = 2$, $A = 1$, and then $g_{n1}, \cdots, g_{nm_n}$ with $S \subset \mathfrak{S}(g_{n1} + (X_n)_i, \cdots, g_{nm_n} + (X_n)_i)$. If an open set T and an $f \epsilon \mathfrak{P}(S, T)$ are given, there is an n for which $\mathfrak{P}(S, f + W_n) \subset \mathfrak{P}(S, T)$, and a ν with $f \epsilon g_{n\nu} + (X_n)_i$. Then $g_{n\nu} + (X_n)_i \subset f - (X_n)_i + (X_n)_i \subset f + (X_n - X_n)_i$ (by Theorem 3) $\subset f + W_n$, and $\mathfrak{P}(S, g_{n\nu} + (X_n)_i) \subset \mathfrak{P}(S, f + W_n) \subset \mathfrak{P}(S, T)$. So in the sequence of open sets $g_{n\nu} + (X_n)_i$, $n = 1, 2, \cdots, \nu = 1, \cdots, m_n$, we can find, for every open set T and every $f \epsilon \mathfrak{P}(S, T)$, a T' with $f \epsilon \mathfrak{P}(S, T') \subset \mathfrak{P}(S, T)$. Therefore they form a complete system of neighborhoods in S.

Ad (8): In separable spaces regularity implies normality (cf. [5]), that

is, (8) follows from (6) and (10).†

7. We are now in a position to formulate and prove the fundamental Theorem 18.

DEFINITION 11. *If D is an arbitrary set, L^D is the set of all functions $F(a)$ with the domain D and a range $\subset L$ (that is, $a \epsilon D$, $F(a) \epsilon L$). A function $F(a)$ is "bounded" if its range (the set of all $F(a)$, $a \epsilon D$) is bounded (see Definition 5, this set being $\subset L$). The set of all bounded $F \epsilon L^D$ is L_b^D. If $U \epsilon \mathfrak{U}$, define a set $U' \subset L_b^D$ as the set of all $F \epsilon L_b^D$ with a range $\subset U$. The set of all U' is $\mathfrak{U}'$.*

Remark. If L is metric, this boundedness means that the (real-numerical) function $\|F(a)\|$ should be bounded. Our $\mathfrak{U}'$ corresponds to the metric $\|F\|'$ of L_b^D defined by $\|F\|' = \text{l.u.b.} \|F(a)\|$.

THEOREM 17. *L_b^D forms with $\mathfrak{U}'$ a topological linear space, that is, it satisfies Definition 2b, (1)–(6). It is convex, that is, it also satisfies Definition 2b, (7), provided that L is convex.*

All parts of this theorem are obvious.

THEOREM 18. *If L is topologically complete, so is L_b^D.*

Remark. It is easily seen that this statement holds for sequential completeness instead of for topological completeness. But, in view of the application to the generalized theory of almost periodic functions, and because we believe that this notion of completeness is the natural one, it is important to prove our theorem in its present form.—

Let $S^* \subset L_b^D$ be a closed and totally bounded set; we have to prove that it is compact. As every infinite $T^* \subset S^*$ contains a sequence $\mathfrak{F}_1, \mathfrak{F}_2, \cdots$ of distinct elements, we may assume $T^* = (\mathfrak{F}_1, \mathfrak{F}_2, \cdots)$.

For a fixed $a \epsilon D$ denote the set of all $\mathfrak{F}(a)$, $\mathfrak{F} \epsilon S^*$, by R_a. As $S^* \subset \mathfrak{S}(\mathfrak{F}_1 + U', \cdots, \mathfrak{F}_n + U')$ implies that $R_a \subset \mathfrak{S}(\mathfrak{F}_1(a) + U, \cdots, \mathfrak{F}_n(a) + U)$, R_a is totally bounded, and, with it, $R_a - R_a$ and $(R_a - R_a)_{\text{cl}}$ (by Theorems 9, 11). The latter set is also closed, and as it is contained in L, it is compact. Thus Theorem 16 applies to it; and, in particular, the open sets $(W_1)_i$, $(W_2)_i$, $\cdots$ constructed in the part "Ad (9)" of its proof form a complete system of neighborhoods of 0 in $(R_a - R_a)_{\text{cl}}$. (Note that the $(W_n)_i$ do not depend on $(R_a - R_a)_{\text{cl}}$, that is, on a; but if we choose for an open set T with $0 \epsilon T$ the $p = p(T, a)$ with $\mathfrak{P}((R_a - R_a)_{\text{cl}}, W_p) \subset \mathfrak{P}((R_a - R_a)_{\text{cl}}, T)$ does depend on T and on a.)

Now apply the diagonal process used in the last paragraph of the proof of Theorem 15 to $\mathfrak{F}_1, \mathfrak{F}_2, \cdots$ and $W_1', W_2', \cdots$ in L_b^D (instead of $f_1, f_2, \cdots$ and

† Tychonoff proves, loc. cit., only normality, that is, Hausdorff's axiom (7) ([1], p. 229). But if we replace in his result ([5], p. 140) the closed sets F, ϕ by two sets F, ϕ without common condensation points, and the space R by $F + \phi$ (in which F, ϕ are relatively closed), (8) obtains. (Hausdorff's notations for F, ϕ, R are F_1, F_2, E.)

$U_1, U_2, \cdots$ in L, which we considered there). As S^* is totally bounded (as S was there) we obtain a subsequence $\mathfrak{F}^{(1)}, \mathfrak{F}^{(2)}, \cdots$, such that, for each W_p', an $n_1' = n_1'(p)$ exists so that $m, n \geqq n_1'$ imply $\mathfrak{F}^{(m)} - \mathfrak{F}^{(n)} \epsilon W_p'$. Thus $\mathfrak{F}^{(m)}(a) - \mathfrak{F}^{(n)}(a) \epsilon W_p$. If any open set $T \subset L$ with $0 \epsilon T$ is given, there is a $p = p(T, a)$ with $\mathfrak{P}((R_a - R_a)_{\mathrm{cl}}, W_p) \subset \mathfrak{P}((R_a - R_a)_{\mathrm{cl}}, T)$. As $\mathfrak{F}^{(m)}(a) - \mathfrak{F}^{(n)}(a) \epsilon R_a - R_a \subset (R_a - R_a)_{\mathrm{cl}}$, we thus have $\mathfrak{F}^{(m)}(a) - \mathfrak{F}^{(n)}(a) \epsilon T$. Putting $n_1(T, a) = n_1'(p(T,a))$, we see that $\mathfrak{F}^{(1)}(a), \mathfrak{F}^{(2)}(a), \cdots$ is a fundamental sequence (in L) in the sense of Definition 9 ($m, n \geqq n_1$ implying $\mathfrak{F}^{(m)}(a) - \mathfrak{F}^{(n)}(a) \epsilon T$). As L is topologically complete, it is also sequentially complete (by Theorem 15), and so $\mathfrak{F}^{(1)}(a), \mathfrak{F}^{(2)}(a), \cdots$ is convergent. Denote its limit by $\mathfrak{F}(a)$ (we know so far only that $\mathfrak{F} \epsilon L^D$).

We constructed above an $n_1' = n_1'(p)$, independent of a, such that, for $m, n \geqq n_1'$, $\mathfrak{F}^{(m)} - \mathfrak{F}^{(n)} \epsilon W_p'$. (Note that, for an arbitrary open set T, $0 \epsilon T$, in the place occupied by W_p, the corresponding $n_1 = n_1(T, a)$ would have depended on a.) Thus $\mathfrak{F}^{(m)}(a) - \mathfrak{F}^{(n)}(a) \epsilon W_p$, $\mathfrak{F}^{(m)}(a) - \mathfrak{F}(a) \epsilon (W_p)_{\mathrm{cl}}$.

Now assume that there exists a $U' \epsilon \mathfrak{U}'$, such that, for infinitely many n, $\mathfrak{F}^{(n)} - \mathfrak{F}$ is not $\epsilon U'$. Then we can select a subsequence $\mathfrak{F}_1^{(1)}, \mathfrak{F}_2^{(1)}, \cdots$ from $\mathfrak{F}^{(1)}, \mathfrak{F}^{(2)}, \cdots$ such that, for all n, $\mathfrak{F}_1^{(n)} - \mathfrak{F}$ is not $\epsilon U'$. Now choose $V \epsilon \mathfrak{U}$, $V + V \subset U$, and repeat the preceding construction with $\mathfrak{F}_1^{(1)}, \mathfrak{F}_2^{(1)}, \cdots$ and $V, W_1, W_2, \cdots$ instead of with $\mathfrak{F}_1, \mathfrak{F}_2, \cdots$ and $W_1, W_2, \cdots$. We obtain a subsequence $\mathfrak{F}_2^{(1)}, \mathfrak{F}_2^{(2)}, \cdots$ of $\mathfrak{F}_1^{(1)}, \mathfrak{F}_1^{(2)}, \cdots$ and a $\mathfrak{G}$ such that, for every $a \epsilon D$, $\mathfrak{G}(a)$ is the limit of $\mathfrak{F}_2^{(1)}(a), \mathfrak{F}_2^{(2)}(a), \cdots$. But this is a subsequence of $\mathfrak{F}^{(1)}(a), \mathfrak{F}^{(2)}(a), \cdots$ and therefore has the limit $\mathfrak{F}(a)$; thus $\mathfrak{F}(a) = \mathfrak{G}(a)$ for all $a \epsilon D$, and $\mathfrak{F} = \mathfrak{G}$. Furthermore, we have, from a certain n on (it is $n_1''(1)$, if $n_1''(p)$ is the analogue of $n_1'(p)$ in the present construction, and thus independent of a) $\mathfrak{F}_2^{(n)}(a) - \mathfrak{F}(a) \epsilon V_{\mathrm{cl}} \subset V + V \subset U$, $\mathfrak{F}_2^{(n)} - \mathfrak{F} \epsilon U'$. But as $\mathfrak{F}_2^{(1)}, \mathfrak{F}_2^{(2)}, \cdots$ is a subsequence of $\mathfrak{F}_1^{(1)}, \mathfrak{F}_1^{(2)}, \cdots$ we should always have $\mathfrak{F}_2^{(n)} - \mathfrak{F}$ is not $\epsilon U'$. This is a contradiction. Thus there is an $n_2' = n_2'(U')$ for every $U' \epsilon \mathfrak{U}'$ such that $n \geqq n_2'$ implies $\mathfrak{F}^{(n)} - \mathfrak{F} \epsilon U'$. (Remember that the sequence $\mathfrak{F}^{(1)}, \mathfrak{F}^{(2)}, \cdots$ is independent of U'.) So if $\mathfrak{F} \epsilon L_b^D$, it is the limit of the sequence $\mathfrak{F}^{(1)}, \mathfrak{F}^{(2)}, \cdots$ and therefore a condensation point of the original sequence $\mathfrak{F}_1, \mathfrak{F}_2, \cdots$. Thus the compactness of S^* would have been established.

Assume $U' \epsilon \mathfrak{U}'$. We proceed as in the last paragraph of the proof of Theorem 8. Choose $V \epsilon \mathfrak{U}$, $V + V \subset U$ and $W \epsilon \mathfrak{U}$, $\alpha W \subset V$ if $-1 \leqq \alpha \leqq 1$. There is an n such that $\mathfrak{F}_n - \mathfrak{F} \epsilon W'$, that is, all $\mathfrak{F}_n(a) - \mathfrak{F}(a) \epsilon W$. The set of all $\mathfrak{F}_n(a)$, $a \epsilon D$, for this n, is bounded, say $\subset \beta W$. Thus, with $\gamma = \max(1, |\beta|)$,

$$\begin{aligned}\mathfrak{F}(a) &= \mathfrak{F}_n(a) - (\mathfrak{F}_n(a) - \mathfrak{F}(a)) \epsilon \beta W - W \\ &= \gamma\left(\frac{\beta}{\gamma} W\right) + \gamma\left(-\frac{1}{\gamma} W\right) \subset \gamma(V + V) \subset \gamma U.\end{aligned}$$

So the set of all $\mathfrak{F}(a)$, $a \epsilon D$, is bounded, and $\mathfrak{F} \epsilon L_b^D$. This completes the proof.

IV. Non-metric examples

8. Many metric and complete linear spaces are known, so that it is not necessary to point out such examples. By Theorem 15, our notion of topological completeness gives nothing new as long as L satisfies Hausdorff's first countability axiom. For this reason those examples will be of particular interest which violate this axiom. We mentioned three such spaces in the last footnote on page 2, and we shall discuss them now in detail.

Definition 12. *Denote Hilbert space by $\mathfrak{H}$ and the space of all bounded linear operators in $\mathfrak{H}$ by $\mathfrak{B}$ (cf. for instance* [6], *Chapter* I; [7], *Chapter* I, *paragraph* 2; [2], *pp.* 372–373). *There are various ways to define sets $\mathfrak{U}$ for $L=\mathfrak{H}$ or $\mathfrak{B}$, of which we shall consider the following.* (*Cf.* [2], *pp.* 378–388, *where the discussion of all these topologies is to be found. The notations $\|f\|$, (f, g), etc. are explained in each of the above references.*)

(a) *For any $\delta>0$ define $U_1(\delta)$ as the set of all $f\epsilon\mathfrak{H}$ with $\|f\| \leqq \delta$; $\mathfrak{U}_1$ is the set of all $U_1(\delta)$.*

(b) *For any $n=1, 2, \cdots, \phi_1, \cdots, \phi_n \epsilon \mathfrak{H}$, $\delta>0$, define $U_2(\phi_1, \cdots, \phi_n; \delta)$ as the set of all $f\epsilon\mathfrak{H}$ with $|(f, \phi_\nu)| \leqq \delta$ for $\nu=1, \cdots, n$; $\mathfrak{U}_2$ is the set of all $U_2(\phi_1, \cdots, \phi_n; \delta)$.*

(c) *For any $\delta>0$ define $U_3(\delta)$ as the set of all $A\epsilon\mathfrak{B}$ with $|||\, A\, ||| \leqq \delta$, where*

$$|||\, A\, ||| = \underset{f\epsilon\mathfrak{H}, f\neq 0}{\text{l.u.b.}} \frac{\|Af\|}{\|f\|}$$

(*that is, the set of all $A\epsilon\mathfrak{B}$ with $\|Af\| \leqq \delta\|f\|$ identically*); *$\mathfrak{U}_3$ is the set of all $U_3(\delta)$.*

(d) *For any $n=1, 2, \cdots, \phi_1, \cdots, \phi_n\epsilon\mathfrak{H}$, $\delta>0$, define $U_4(\phi_1, \cdots, \phi_n; \delta)$ as the set of all $A\epsilon\mathfrak{B}$ with $\|A\phi_\nu\| \leqq \delta$ for $\nu=1, \cdots, n$; $\mathfrak{U}_4$ is the set of all $U_4(\phi_1, \cdots, \phi_n; \delta)$.*

(e) *For any $n=1, 2, \cdots, \phi_1, \psi_1, \cdots, \phi_n, \psi_n\epsilon\mathfrak{H}$, $\delta>0$, define $U_5(\phi_1, \psi_1, \cdots, \phi_n, \psi_n; \delta)$ as the set of all $A\epsilon\mathfrak{B}$ with $|(A\phi_\nu, \psi_\nu)| \leqq \delta$ for $\nu=1, \cdots, n$; $\mathfrak{U}_5$ is the set of all $U_5(\phi_1, \psi_1, \cdots, \phi_n, \psi_n; \delta)$.*

$\mathfrak{U}_1$ describes the strong, $\mathfrak{U}_2$ the weak topology of $\mathfrak{H}$; $\mathfrak{U}_3$ describes the uniform, $\mathfrak{U}_4$ the strong, $\mathfrak{U}_5$ the weak topology of $\mathfrak{B}$.

Theorem 19. *$\mathfrak{H}$ and $\mathfrak{B}$ are convex topological linear spaces* (*that is, they fulfill Definition* 2b, (1)–(7)) *in all five topologies of Definition* 11.

The proof is immediate.

THEOREM 20. *$\mathfrak{H}$, $\mathfrak{U}_1$ and $\mathfrak{B}$, $\mathfrak{U}_3$ are originated by metrics (in the sense of the remark after Definition* 2b), *with the absolute values $\|f\|$ and $|||A|||$ respectively. Both are sequentially, and thus topologically, complete.*

The metric properties are verified immediately. Sequential completeness is one of the fundamental properties of $\mathfrak{H}$, $\mathfrak{U}_1$ (cf. [6], pp. 66 and 111), and extends from it immediately to $\mathfrak{B}$, $\mathfrak{U}_3$. Topological completeness follows by Theorem 15.

We now investigate the non-metric topologies $\mathfrak{H}$, $\mathfrak{U}_2$ and $\mathfrak{B}$, $\mathfrak{U}_4$ or $\mathfrak{U}_5$. Theorem 22 has some independent interest.

THEOREM 21. *A set $S\subset\mathfrak{H}$ or $\mathfrak{B}$ is totally bounded in the topology $\mathfrak{U}_2$ or $\mathfrak{U}_5$ (these are the weak topologies), if and only if the (real-numerical) sets of all $|(f,\phi)|$, $f\epsilon S$, or $|(A\phi,\psi)|$, $A\epsilon S$, are bounded for every choice of $\phi\epsilon\mathfrak{H}$ or ϕ, $\psi\epsilon\mathfrak{H}$ (no uniformity is required).*

Consider $\mathfrak{H}$, $\mathfrak{U}_2$. If S is totally bounded, take $U=U_2(\phi;\ 1)$, $S\subset\mathfrak{S}(f_1+U,\cdots,f_n+U)$. Then each $f\epsilon S$ is such that $|(f-f_\nu,\phi)|\leqq 1$ for some $\nu=1,\cdots,n$, and thus

$$|(f,\phi)| = \max_{\nu=1,\cdots,n}|(f_\nu,\phi)| + 1,$$

proving the necessity of our condition. To prove its sufficiency, consider a $U\epsilon\mathfrak{U}_2$, that is, $U=U_2(\phi_1,\cdots,\phi_n;\ \delta)$. Choose $c\geqq|(f,\phi_\nu)|$ for all $f\epsilon S$, $\nu=1,\cdots,n$, and choose $N=1,2,\cdots$ with $2^{3/2}c/N\leqq\delta$. Consider all N^{2n} systems of $2n$ integers $\rho_1,\sigma_1,\cdots,\rho_n,\sigma_n=-N+1,-N+3,\cdots,N-1$. If for a combination $\rho_1,\sigma_1,\cdots,\rho_n,\sigma_n$ there exists an f with

$$\frac{\rho_\nu-1}{N}c \leqq \Re(f,\phi_\nu) \leqq \frac{\rho_\nu+1}{N}c,$$

$$\frac{\sigma_\nu-1}{N}c \leqq \Im(f,\phi_\nu) \leqq \frac{\sigma_\nu+1}{N}c \quad \text{for all } \nu=1,\cdots,n,$$

choose one and call it $f_{\rho_1\sigma_1\cdots\rho_n\sigma_n}$. The number of these $f_{\rho_1\sigma_1\cdots\rho_n\sigma_n}$ is finite, say $M\leqq N^{2n}$; let us denote them by $f^{(1)},\cdots,f^{(M)}$. One easily sees that, for each $f\epsilon S$, there exists some $f^{(\mu)}$ with

$$|\Re(f-f^{(\mu)},\phi_\nu)| \leqq \frac{2c}{N},\ |\Im(f-f^{(\mu)},\phi_\nu)| \leqq \frac{2c}{N} \text{ for all } \nu=1,\cdots,n,$$

implying that

$$|(f-f^{(\mu)},\phi_\nu)| \leqq \frac{2^{3/2}c}{N} \leqq \delta.$$

Thus $S \subset \mathfrak{S}(f^{(1)}+U, \cdots, f^{(M)}+U)$, proving the total boundedness of our set S, and the sufficiency of our condition.

$\mathfrak{B}$, $\mathfrak{U}_5$ are discussed in the same way.

THEOREM 22. *A set $S \subset \mathfrak{H}$ or $\mathfrak{B}$ is totally bounded in the topology $\mathfrak{U}_2$ or $\mathfrak{U}_5$, if and only if it is bounded in the topology $\mathfrak{U}_1$ or $\mathfrak{U}_3$ (these are the metric topologies), that is, if the (real-numerical) set of all $\|f\|$, $f \epsilon S$, or $|\|A\||$, $A \epsilon S$, is bounded.*

Consider $\mathfrak{H}$, $\mathfrak{U}_2$. If $\|f\| \leqq c$, then $|(f, \phi)| \leqq \|f\| \cdot \|\phi\| \leqq c\|\phi\|$, proving the total boundedness by Theorem 21, and thus the sufficiency of our condition. To prove the necessity, assume S totally bounded and the absolute values $\|f\|$, $f \epsilon S$, not bounded. For each $n=1, 2, \cdots$ choose $f_n \epsilon S$, $\|f_n\| \geqq n^2$. For each ϕ, $|(f_n, \phi)|$ are bounded (by Theorem 21). Thus $\lim_{n\to\infty} \|f_n/n\| = \infty$ and, for each ϕ, $\lim_{n\to\infty} (f_n/n, \phi)=0$, contradicting a basic property of "weak convergence" (cf. [2], p. 380, footnote 32, those considerations being a special case of a result of S. Banach, cf. [8], Theorem 5, pp. 157–160). Thus our condition is necessary.

$\mathfrak{B}$, $\mathfrak{U}_5$ are discussed in the same way (using [2], p. 382, footnote 35).

THEOREM 23. *Each of the three topologies $\mathfrak{U}_2$, for $\mathfrak{H}$, $\mathfrak{U}_4$ and $\mathfrak{U}_5$ for $\mathfrak{B}$ violate both countability axioms of Hausdorff, but all three are topologically complete.*

Hausdorff's first countability axiom is not fulfilled, by [2], p. 380 and pp. 382–383; hence the second is not fulfilled either. Every $\mathfrak{U}_2$-totally bounded set $S \subset \mathfrak{H}$ is a subset of some set $U_1(c)$: $\|f\| \leqq c$. Now the latter sets are all compact (cf. [2], p. 381, footnote 34), and S, being a closed subset of a compact set, is also compact. Therefore $\mathfrak{H}$ with $\mathfrak{U}_2$ is topologically complete. $\mathfrak{B}$ with $\mathfrak{U}_5$ is discussed in the same way.

It remains to discuss $\mathfrak{B}$ with $\mathfrak{U}_4$. We may replace every $A \epsilon \mathfrak{B}$ by its adjoint A^*; then $A \epsilon \mathfrak{U}_4(\phi_1, \cdots, \phi_n, \delta)$ means that $\|A^*\phi_\nu\| \leqq \delta$ for $\nu=1, \cdots, n$. This means that, for every $f \epsilon \mathfrak{H}$ with $\|f\|=1$, $|(A^*\phi_\nu, f)| \leqq \delta$ (the sufficiency follows from Schwarz's inequality, the necessity from the substitution $f=A^*\phi_\nu/\|A^*\phi_\nu\|$), that is, $|(Af, \phi_\nu)| \leqq \delta$. The bounded linear operators A are characterized by the properties $A(\alpha f)=\alpha Af$ (α any complex number), $A(f+g)=Af+Ag$ within the set $\mathfrak{B}_1$ of all operators A, subject only to the restriction $A(\alpha f)=\alpha Af (\alpha>0)$. The boundedness condition $|(Af, g)| \leqq c \cdot \|f\| \cdot \|g\|$ (or $\|Af\| \leqq c \cdot \|f\|$, cf. the remark made above), with some fixed $c>0$, is required in any case. If we use the topology analogous to $\mathfrak{U}_4$ in $\mathfrak{B}_1$, $\mathfrak{B}$ is a closed subset of $\mathfrak{B}_1$, therefore topologically complete if $\mathfrak{B}_1$ is topologically complete. For the operators A of $\mathfrak{B}_1$ we may restrict the definition domain to the set S_1: $\|f\|=1$, as $A(\alpha f)=\alpha Af(\alpha>0)$ then allows a unique extension to $\mathfrak{H}$. Now in this interpretation $\mathfrak{B}_1$, $\mathfrak{U}_4$ coincides exactly with the $\mathfrak{H}_b^{S_1}$ of $\mathfrak{H}$, $\mathfrak{U}_2$, from

Definition 11, and thus it is topologically complete by Theorem 18. This completes the proof.

V. THE PSEUDO-METRICS IN CONVEX SPACES

9. If L is topological and convex, that is, if L fulfills Definition 2b, (1)–(7), we can define a family of notions, each of which is an analogue to the absolute value, and which together describe the topology of L. In various applications (for instance, in the theory of almost periodic functions) this can be used to replace the metric, even when the countability axioms are violated.

THEOREM 24. *If L is topological and convex, $(\alpha U)_{cl} \subset (\beta U)_i$ for $U \epsilon \mathfrak{U}$ and $0<\alpha<\beta$.*

By Definition 2b, (7), $U+U \subset 2U$; repeating this n times, $U+ \cdots +U$ (2^n times) $\subset 2^nU$, and thus, a fortiori, $(2^n-2)U+U+U \subset 2^nU$. Therefore $(2^n-2)U_{cl} \subset (2^n-2)U+U$ (by Theorem 5) $\subset (2^nU)_i = 2^nU_i$, $((2^n-2)/2^n)U_{cl} \subset U_i$. By Theorem 12, $0<\alpha<1$ implies $\alpha U_{cl} \subset \alpha U_{cl}+(1-\alpha)U_{cl} = U_{cl}$, and, upon replacing α by α/β and multiplying by β, $0<\alpha<\beta$, $\alpha U_{cl} \subset \beta U_{cl}$. Now for $0<\alpha<\beta$ we can find an integer n for which $\alpha/\beta<(2^n-2)/2^n<1$, and then

$$\alpha U_{cl} \subset \beta\left(\frac{2^n-2}{2^n} U_{cl}\right) \subset \beta U_i.$$

DEFINITION 13. *For $f \epsilon L$ and $U \epsilon \mathfrak{U}$ consider the set of all $\alpha>0$ with $f \epsilon \alpha U$. Its* g.l.b. *is denoted by $\|f\|_U^+$.*

THEOREM 25. *$\|f\|_U^+$ is finite, $\geqq 0$, and continuous. If $\alpha \geqq 0$, $\|\alpha f\|_U^+ = \alpha \|f\|_U^+$; furthermore, $\|f+g\|_U^+ \leqq \|f\|_U^+ + \|g\|_U^+$. $\|f\|_U^+=0$ means that $f \epsilon \mathfrak{P}$ (αU over all $\alpha>0$) $=(+0)U$.*†

As βf is continuous, for small $\beta>0$, $\beta f \epsilon U$, $f \epsilon U/\beta$; thus the α-set is not empty and $\|f\|_U^+$ is finite; it is obviously non-negative. $\|\alpha f\|_U^+ = \alpha\|f\|_U^+$ is obvious for $\alpha>0$; for $\alpha=0$ it states merely that $\|0\|_U^+=0$. If $f \epsilon \alpha U$, $g \epsilon \beta U$, Theorems 12 and 24 imply that $f+g \epsilon (\alpha+\beta)U_{cl} \subset (\alpha+\beta+\delta)U$ for every $\delta>0$; from this it follows that $\|f+g\|_U^+ \leqq \|f\|_U^+ + \|g\|_U^+$. The last statement is obvious. It remains to prove the continuity of $\|f\|_U^+$.

Assume $0<\alpha<\|f_0\|_U^+<\beta$. (If $\|f_0\|_U^+=0$ then omit the α.) Choose α', β' with $\alpha<\alpha'<\|f_0\|_U^+<\beta'<\beta$. Then $f_0 \epsilon \beta' U \subset \beta U_i$, f_0 is not $\epsilon \alpha' U \supset \alpha U_{cl}$, $f_0 \epsilon \mathfrak{P}(\beta U_i, \mathfrak{C}(\alpha U_{cl}))$. If an f belongs to this set, we have $f \epsilon \beta U_i \subset \beta U$, f is not ϵ $\alpha U_{cl} \supset \alpha U$, and thus $\alpha \leqq \|f\|_U^+ \leqq \beta$. Furthermore, $\mathfrak{P}(\beta U_i, \mathfrak{C}(\alpha U_{cl}))$ is obviously open, proving the continuity of $\|f\|_U^+$.

---

† We write $(+0)U$ in order to distinguish this set from $0U$, which of course is (0).

THEOREM 26. *The sets* $\|f-f_0\|_U^+<\delta$, $f_0\epsilon L$, $U\epsilon\mathfrak{U}$, $\delta>0$, *form a complete system of neighborhoods in* L. (*It would be sufficient to consider* $\delta=1$.)

As $\|f-f_0\|_U^+$ is continuous in f, these sets are all open. If S is an open set, $f_0\epsilon S$, there is a $U\epsilon\mathfrak{U}$ with $f_0+U\subset S$, and $\|f-f_0\|_U^+<1$ implies $f-f_0\epsilon U$, $f\epsilon f_0+U\subset S$. Thus the set $\|f-f_0\|_U^+<1$ contains f_0 and is contained in S.

We could extend $\|\alpha f\|_U^+=\alpha\|f\|_U^+$ from the non-negative α's to all real α by introducing $\|f\|_U=\max(\|f\|_U^+, \|-f\|_U^+)$, but we shall not discuss this further here.

Note that in metric spaces L (cf. the remark after Definition 2b), the functions $\|f\|_U^+$ with $U=S^o(0;\delta)$ or with $U=S^{cl}(0;\delta)$ coincide with each other and with their $\|f\|_U$, all of them being equal to $\|f\|/\delta$. If L is only topological and if we form L_b^D by Definition 11, then obviously

$$\|\mathfrak{F}\|_{U'}^+ = \underset{x\epsilon D}{\mathrm{l.u.b.}}\ \|\mathfrak{F}(x)\|_U^+.$$

APPENDIX I

10. We wish to make two remarks which are useful for some applications.

Remark 1. The linear space L, as defined in Definition 1, may allow complex numbers α (instead of the real ones alone) as factors. We then call L complex linear and we change Definition 1 so as to admit in its conditions (4), (5), (6) also complex α and β. Then Definition 2a for the metric should be formulated with complex α's in its condition (2). But the important change is the one in Definition 2b for topological L's: here condition (4) must include all complex α's with $|\alpha|\leqq 1$ instead of only the real α's with $-1\leqq\alpha\leqq 1$. Then Theorem 7, stating the continuity of αf as a function of α and f, can be proved without any changes. (Note that the definition of convexity, Definition 2b, (7), is unaffected, and that Definition 7, Theorems 12 and 13 remain restricted to real coefficients.)

This stronger form of Definitions 2b, (4), which we call (4′), can be replaced by the following two conditions:

($4_1'$) if $U\epsilon\mathfrak{U}$, there is a $V\epsilon\mathfrak{U}$, such that, for every real α with $0\leqq\alpha\leqq 1$, $\alpha V\subset U$,

($4_2'$) if $U\epsilon\mathfrak{U}$, there is a $V\epsilon\mathfrak{U}$ with $iV\subset U$.

($4_1'$) and ($4_2'$), with the help of (3) and (5), include (4′): Choose $V+V\subset U$, W with $\alpha W\subset V$ for $0\leqq\alpha\leqq 1$, X with $iX\subset W$, Y with $iY\subset X$, Z with $iZ\subset Y$, $\Xi\subset\mathfrak{P}(W, X, Y, Z)$, where V, W, X, Y, Z, Ξ are all $\epsilon\mathfrak{U}$. Thus $i^n\Xi\subset W$ for $n=0, 1, 2, 3$. Now if a complex α is such that $|\alpha|\leqq 1$, we can write $\alpha=i^m\beta+i^n\gamma$, $m=0, 2$; $n=1, 3$; β, γ real; $0\leqq\beta, \gamma\leqq 1$. Thus $\alpha\Xi\subset\beta i^m\Xi+\gamma i^n\Xi\subset\beta W+\gamma W\subset V+V\subset U$, proving (4′) as we stated it.

Remark 2. If L is convex and $S \subset L$, then

$$((S_{\mathrm{conv}})_{\mathrm{cl}})_{\mathrm{conv}} = (S_{\mathrm{conv}})_{\mathrm{cl}}.$$

As the left side obviously $\supset$ the right side, we need only prove $\subset$. And as $(S_{\mathrm{conv}})_{\mathrm{conv}} = S_{\mathrm{conv}}$, we may write T for S_{conv} and prove $(T_{\mathrm{cl}})_{\mathrm{conv}} \subset (T_{\mathrm{conv}})_{\mathrm{cl}}$. Assume $U \epsilon \mathfrak{U}$, choose $V \epsilon \mathfrak{U}$, $V+V \subset U$. Then $V_{\mathrm{conv}} \subset U$ (cf. the beginning of the proof of Theorem 14), and $T_{\mathrm{cl}} \subset T+V$, $(T_{\mathrm{cl}})_{\mathrm{conv}} \subset T_{\mathrm{conv}}+V_{\mathrm{conv}} = T_{\mathrm{conv}}+U$. Thus $(T_{\mathrm{cl}})_{\mathrm{conv}} \subset \mathfrak{P}(T_{\mathrm{conv}}+U)$, where U runs over all elements of $\mathfrak{U}$. By Theorem 5, the right side is $=(T_{\mathrm{conv}})_{\mathrm{cl}}$, thus $(T_{\mathrm{cl}})_{\mathrm{conv}} \subset (T_{\mathrm{conv}})_{\mathrm{cl}}$, completing the proof.

Appendix II

11. The coefficients α, which occur in Definition 1, play an important role in the applications of this theory, but the theory itself could be developed without them. That is, we could work on the basis of Definition 1, parts (1), (2), (7) alone; in other words, the theory could be extended from linear spaces to Abelian groups (except, of course, for the statements about convexity). Definition 2a has then to be restricted to (1), (3), and Definition 2b to (1), (2), (3), (4) with $\alpha = -1$ only, and (5). Note that (1), (2), (3) are general topological axioms, while (4), (5) express that $-f$ and $f+g$ (that is, the "group operations") are continuous.

After these changes are made all our considerations remain almost unaltered, except that the notion of "boundedness" must be avoided (because Definition 2b, (6), on which it is based, has been omitted); it has to be replaced by "total boundedness."

Bibliography to "Complete topological spaces"

1. F. Hausdorff, *Mengenlehre*, Berlin, deGruyter, 1927.

2. J. von Neumann, *Zur Algebra der Funktionaloperatoren und Theorie der normalen Operatoren*, Mathematische Annalen, vol. 102 (1929), pp. 370–427.

3. P. Alexandroff and P. Urysohn, *Zur Theorie der topologischen Räume*, Mathematische Annalen, vol. 92 (1924), pp. 258–266.

4. P. Alexandroff, *Über die Struktur der bikompakten topologischen Räume*, Mathematische Annalen, vol. 92 (1924), pp. 267–274.

5. A. Tychonoff, *Über einen Metrisationssatz von P. Urysohn*, Mathematische Annalen, vol. 95 (1926), pp. 139–142.

6. J. von Neumann, *Allgemeine Eigenwerttheorie Hermitescher Funktionaloperatoren*, Mathematische Annalen, vol. 102 (1929), pp. 49–131.

7. M. H. Stone, *Linear Transformations in Hilbert Space*, American Mathematical Society Colloquium Publications, vol. 15, New York, 1932.

8. S. Banach, *Sur les opérations dans les ensembles linéaires et leur application aux équations intégrales*, Fundamenta Mathematicae, vol. 3 (1922), pp. 133–181.

Institute for Advanced Study,
Princeton, N. J.

ALMOST PERIODIC FUNCTIONS IN GROUPS, II*

BY

S. BOCHNER AND J. VON NEUMANN

The present paper is a continuation of the article by J. von Neumann on *Almost periodic functions in a group*, I [1].† Its main object is to extend the theory of almost periodicity to those functions having values which are not numbers but elements of a general linear space L. For functions of a real variable this extension was begun by Bochner [2], and then applied by him, see [3], to a problem concerning partial differential equations.

Bochner assumed L to be both complete and metric. In the present paper we shall admit more general linear spaces. We shall drop the metric but keep the completeness. Since the usual notion of completeness is based on the notion of metric, it was necessary to establish, for linear spaces, a notion of completeness independent of it. This was done in the preceding note of J. von Neumann [4]. The results of this note will be employed throughout, and we observe that, from the very beginning, we shall assume that L is linear with respect to arbitrary complex coefficients, see [4], Appendix I.

As in [1], the main difficulty to overcome was the definition and the establishment of a mean. This was done in Part I. The definition of a mean remained actually the same as in [1], but the proof of the existence of a mean necessitated a more elaborate argument, although, in broad lines, the argument does not differ essentially.

In Part II we deduce the existence and uniqueness of a Fourier expansion for any almost periodic function. It is worth pointing out that the representations occurring in the Fourier expansions of abstract almost periodic functions are the same as for numerical almost periodic functions, only the constant coefficients by which the representations are multiplied are abstract elements instead of numbers. (More than that, if in a linear manifold L different topologies are suitable for our purposes, then even the nature of the coefficients no longer determines the precise nature of abstractness of the almost periodic function.) Thus, roughly speaking, there are no more abstract almost periodic functions than numerical almost periodic functions. In particular, if a group admits of no other numerical almost periodic functions than the constant ones, there exists no non-constant abstract almost periodic function, no matter how general the range-space L may be.

---

* Presented to the Society, December 28, 1934; received by the editors June 7, 1934.

† Numbers in brackets refer to the bibliography at the end of this paper.

In Part III we deduce the approximation theorem. Moreover, what is new also for numerical functions, we deduce summation theorems, that is to say, theorems concerning the construction of an almost periodic function of which only the Fourier expansion is given. For functions of Bohr these theorems were established by Bochner [5].

Finally, in Part IV we consider the special case in which L is a Hilbert space, or the space of bounded linear transformations of a Hilbert space into itself. We particularly refer to Theorems 39 and 40. Theorem 39 treats one of the rare cases of a class of almost periodic functions for which the Fourier expansions may be completely characterized by direct properties. Theorem 40 implies a necessary and sufficient criterion for a locally compact separable group to be compact. We should also mention that for the Abelian addition group of all integers, R. H. Cameron, in an unpublished paper, has found a result which has some connection with our Theorem 39.

PART I. THE MEAN-VALUE

Let L be a convex topological space which we shall assume throughout to be topologically complete (cf. [4], Definitions 1, 2b, and 10); and let $\mathfrak{G}$ denote a fixed arbitrary group. The elements of $\mathfrak{G}$ will be denoted by $a, b, c, \cdots, x, y, z, \cdots$. *Whenever a function is not specified, it will be tacitly assumed to be an element of* $L_b^{\mathfrak{G}}$. ($L_b^{\mathfrak{G}}$ is the set of all bounded functions with the domain $\mathfrak{G}$ and a range $\subset L$, cf. [4], Definition 11. $L_b^{\mathfrak{G}}$ is a topologically complete convex topological set, cf. [4], Theorem 18.)

If $F \epsilon L_b^{\mathfrak{G}}$, the "translated" functions $r_aF = F(xa)$, $l_aF = F(ax)$ also belong to $L_b^{\mathfrak{G}}$ for any $a \epsilon \mathfrak{G}$. We shall denote by $\mathfrak{R}_F$ the set of all r_aF, and by $\mathfrak{L}_F$ the set of all $l_aF(a \epsilon \mathfrak{G})$.

DEFINITION 1. *A function F is almost periodic if both sets $\mathfrak{R}_F$ and $\mathfrak{L}_F$ are totally bounded* (cf. [4], Definition 6). *The set of all almost periodic functions will be denoted by Ap.*

THEOREM 1. *Every constant function is almost periodic.*

The proof is obvious.

DEFINITION 2. *Given groups $\mathfrak{G}_1, \mathfrak{G}_2, \cdots, \mathfrak{G}_p$; we denote by $\mathfrak{G}_1 \times \mathfrak{G}_2 \times \cdots \times \mathfrak{G}_p$ the group consisting of all p-tuples*

$$x = [x_1, \cdots, x_p] \qquad (x_1 \epsilon \mathfrak{G}_1, \cdots, x_p \epsilon \mathfrak{G}_p)$$

with the rules

$$[x_1, \cdots, x_p][y_1, \cdots, y_p] = [x_1y_1, \cdots, x_py_p],$$
$$[x_1, \cdots, x_p]^{-1} = [x_1^{-1}, \cdots, x_p^{-1}].$$

And we shall say that a function is "almost periodic in $x_1, \cdots, x_p$" *if it is almost periodic in* $[x_1, \cdots, x_p]$ *in the group* $\mathfrak{G}_1 \times \mathfrak{G}_2 \times \cdots \times \mathfrak{G}_p$. *In case* $\mathfrak{G}_1 = \mathfrak{G}_2 = \cdots = \mathfrak{G}_p = \mathfrak{G}$, *we shall write* $\mathfrak{G}^p$ *for* $\mathfrak{G}_1 \times \mathfrak{G}_2 \times \cdots \times \mathfrak{G}_p$.

THEOREM 2. *If* F *is almost periodic, the set of* (x, y)*-functions* $F_a = F(xay)$ $(\epsilon L^{\mathfrak{G}^2})$ *is totally bounded* $(a \epsilon \mathfrak{G})$.

Given $U \epsilon \mathfrak{U}$, we choose $V \epsilon \mathfrak{U}$, with $V+V-V \epsilon U$. We choose elements $b_1, \cdots, b_n$ such that to any y there corresponds an index $\nu = \nu(y)$ for which

$$F(zy) - F(zb_\nu) \epsilon V \qquad (z \epsilon \mathfrak{G}).$$

(Remember for this, and for all discussions below, [4], Definition 6.) Hence

$$F(xay) - F(xab_\nu) \epsilon V \qquad (x, a \epsilon \mathfrak{G}).$$

We consider the n functions $F_\nu(x) = F(xb_\nu)$. For each ν the set of x-functions $F_\nu(xa)$ is totally bounded in a. By a simple argument (compare [1], the corresponding part in the proof of Theorem 9) it follows that there exist elements $a_1, \cdots, a_m$, and to each a there corresponds a $\mu = \mu(a)$ such that

$$F_\nu(xa) - F_\nu(xa_\mu) \epsilon V \qquad (x \epsilon \mathfrak{G};\ \nu = 1, \cdots, n).$$

Hence to each a there corresponds a $\mu = \mu(a)$ such that, for all $x, y \epsilon \mathfrak{G}$,

$$\begin{aligned} F(xay) - F(xa_\mu y) &= (F(xay) - F(xab_\nu)) + (F(xab_\nu) - F(xa_\mu b_\nu)) \\ &\quad + (F(xa_\mu b_\nu) - F(xa_\mu b)) \epsilon V + V - V \subset U. \end{aligned}$$

THEOREM 3. *Let* $F(z)$ *be an almost periodic function. Let* z *be the product of positive or negative integer powers of elements* $x', x'', \cdots, x^{(p)}, a', a'', \cdots, a^{(q)} \epsilon \mathfrak{G}$ *in an arbitrary fixed order, and let* $F(z)$ *be considered as a function*

$$(1) \qquad F_{a', \cdots, a^{(q)}} = F_{a', \cdots, a^{(q)}}(x', \cdots, x^{(p)})$$

of $L^{\mathfrak{G}^p}$, *depending on the parameters* $a', \cdots, a^{(q)}$. *The set of these functions is totally bounded in* $a', \cdots, a^{(q)}$.

It is easily seen that in this proof we may replace in z any set of consecutive factors which are all powers of variables or all powers of parameters by a new variable or a new parameter respectively (this may increase the indices p and q). Thus we may assume that z has the form $a'x'a''x'' \cdots$ or $x'a'x''a'' \cdots$. We denote the number of factors in z by k. For $k=2$, our theorem holds by Definition 1, and we are going to apply induction from k to $k+1$. Denoting the product of the first $k-1$ factors in z by ξ, we have to dispose of two cases: (i) $z = \xi xa$, (ii) $z = \xi ax$. A subscript to ξ shall indicate that special values have been assigned to all parameters occurring in ξ.

In Case i we know that to each $U \epsilon \mathfrak{U}$ there correspond quantities $\xi_1, \cdots, \xi_n$ with

(2) $$F(\xi x) \in \mathop{\mathfrak{S}}_{\nu=1}^{n} (F(\xi_\nu x) + V').$$

(Cf. [4], Definitions 6, 11.) Given $U \in \mathfrak{A}$, let $V + V \subset U$. Replacing x by xa in (2) we get

(3) $$F(\xi xa) \in \mathop{\mathfrak{S}}_{\nu=1}^{n} (F(\xi_\nu xa) + V').$$

We determine elements $a_1, \cdots, a_m$ such that

(4) $$F(za) \in \mathop{\mathfrak{S}}_{\mu=1}^{m} (F(za_\mu) + V').$$

Putting here $z = \xi_\nu x$ and substituting the result in (3), we obtain

$$F(\xi xa) \in \mathop{\mathfrak{S}}_{\mu=1}^{m} \mathop{\mathfrak{S}}_{\nu=1}^{n} (F(\xi_\nu xa_\mu) + V' + V') \subset \mathop{\mathfrak{S}}_{\mu=1}^{m} \mathop{\mathfrak{S}}_{\nu=1}^{n} (F(\xi_\nu xa_\mu) + U').$$

This proves Case i.

In Case ii we determine quantities $\xi_1, \cdots, \xi_n$, such that

$$F(\xi z) \in \mathop{\mathfrak{S}}_{\nu=1}^{n} (F(\xi_\nu z) + V');$$

hence

(5) $$F(\xi ax) \in \mathop{\mathfrak{S}}_{\nu=1}^{n} (F(\xi_\nu ax) + V').$$

By Theorem 2 we may determine elements $a_1, \cdots, a_m$ such that

$$F(yax) \in \mathop{\mathfrak{S}}_{\mu=1}^{m} (F(ya_\mu x) + V').$$

Putting here $y = \xi_\nu$ and substituting the result in (5), we obtain

$$F(\xi ax) \in \mathop{\mathfrak{S}}_{\mu=1}^{m} \mathop{\mathfrak{S}}_{\nu=1}^{n} (F(\xi_\nu a_\mu x) + U').$$

This proves Case ii.

COROLLARY. *Let $F(z)$ be an almost periodic function. Let z be the product of positive or negative integer powers of variables $x', x'', \cdots, x^{(p)}$, of parameters $a', a'', \cdots, a^{(q)}$, and of constant elements $c', c'', \cdots, c^{(r)}$, in an arbitrary fixed order. The set of functions $F(z)$ $(\epsilon L_b^{\mathfrak{G}^p})$ is again totally bounded in $a', a'', \cdots, a^{(q)}$.*

If we consider also the elements $c', \cdots, c^{(r)}$ as parameters, we get a larger set of functions $\epsilon L^{\mathfrak{G}^r}$ which by Theorem 3 is totally bounded. But a subset of a totally bounded set is also totally bounded.

THEOREM 4. *The function* (1) *of Theorem* 3 *is almost periodic in* $x', \cdots, x^{(p)}$, *for any fixed values of the parameters* $a', \cdots, a^{(q)}$.

If we multiply the element $[x', \cdots, x^{(p)}] \epsilon \mathfrak{G}^p$ by an element $[b', \cdots, b^{(p)}]$ on the right or the left, the argument z in $F(z)$ goes over into a product of positive or negative integer powers of the variables x, the constants a, and the parameters b. Replacing in the corollary of Theorem 3 the letters a and c by b and a, we find that the resulting set of functions is totally bounded in the parameters b. By Definition 1, the function (1) is almost periodic.

THEOREM 5. *If* $F_1, \cdots, F_k$ *are almost periodic functions with values in linear spaces* $L_1, \cdots, L_k$ *respectively, if* S_κ $(\kappa=1, \cdots, k)$ *is the range of* F_κ, *and if* $\mathfrak{F}(f_1, \cdots, f_k)$ *is a function with the domain* $f_\kappa \epsilon S_\kappa$, $\kappa=1, \cdots, k$, *uniformly continuous in it, and with a range* $\subset L$, *then the function*

$$G(x) = \mathfrak{F}(F_1(x), \cdots, F_k(x))$$

is almost periodic.

Definition 1 is fulfilled for the function $G(x)$ on account of [4], Theorem 9, if this theorem is applied to the function $\mathfrak{F}(F_1, \cdots, F_k)$, considered as a function of the functions $F_1, \cdots, F_k$, with the L_κ, L of Theorem 9 in [4] put equal to $L_{\kappa b}^{\mathfrak{G}}$, $L_b^{\mathfrak{G}}$, and its ranges S_κ to the $\mathfrak{R}_{F_\kappa}$, $\mathfrak{L}_{F_\kappa}$ respectively $(\kappa=1, \cdots, k)$.

THEOREM 6. *If* $F, G \epsilon Ap$, *then* $F \pm G \epsilon Ap$. *If* $F \epsilon Ap$, *and* $\alpha(x)$ *is a numerical almost periodic function, then* $\alpha F \epsilon Ap$.

If $F \epsilon Ap$, the set of functions $r_a F = F(xa)$ is totally bounded; if we put $x=1$ we find that the range of F is totally bounded. Theorem 6 follows from Theorem 5, since the functions f_1+f_2, $f_1 f_2$ are uniformly continuous if f_1 and f_2 run over totally bounded sets, the closure of such sets being compact and separable ([4], Theorem 11, Theorem 7, Definition 10, and Theorem 16).

THEOREM 7. *The set* Ap *is closed.*

Let F be a condensation point of Ap, $U \epsilon \mathfrak{U}$, $V+V \subset U$. There is a $G \epsilon Ap$ such that $F \epsilon G + V'$. Obviously $r_a F \epsilon r_a G + V'$. Choose elements $F_1, \cdots, F_n \epsilon Ap$, such that

$$\mathfrak{R}_G \subset \mathop{\mathfrak{S}}_{\nu=1}^{n} (F_\nu + V').$$

Then

$$\mathfrak{R}_F \subset \mathop{\mathfrak{S}}_{\nu=1}^{n} (F_\nu + V' + V') \subset \mathop{\mathfrak{S}}_{\nu=1}^{n} (F_\nu + U'),$$

which proves that $\mathfrak{R}_F$ is totally bounded. Similarly, $\mathfrak{L}_F$ is totally bounded.

COROLLARY. *If a sequence of almost periodic functions is uniformly convergent, the limit function is also almost periodic.*

THEOREM 8. *If $F \epsilon Ap$, the set $S=((\mathfrak{R}_F)_{\text{conv}})_{\text{cl}}$ has the following properties:* (i) *$S \subset Ap$,* (ii) *S is compact and separable,* (iii) *$S_{\text{conv}}=S$. The same is true also for $T=((\mathfrak{L}_F)_{\text{conv}})_{\text{cl}}$.*

(i) follows from Theorems 4, 6 and 7. (ii) follows from the fact that S is totally bounded ([4], Theorems 14 and 16). (iii) was proved in [4] Appendix I.

THEOREM 9. *If $G \epsilon ((\mathfrak{R}_F)_{\text{conv}})_{\text{cl}}$, then*

$$((\mathfrak{R}_G)_{\text{conv}})_{\text{cl}} \subset ((\mathfrak{R}_F)_{\text{conv}})_{\text{cl}}. \tag{6}$$

The same is true if we replace $\mathfrak{R}$ by $\mathfrak{L}$.

For any $U \epsilon \mathfrak{U}$, we have by assumption ([4], Theorem 5): $G \subset (\mathfrak{R}_F)_{\text{conv}}+U'$. This immediately leads to the result that $\mathfrak{R}_G \subset (\mathfrak{R}_F)_{\text{conv}}+U'$. Since U is arbitrary, it follows ([4], Theorem 5) that $\mathfrak{R}_G \subset ((\mathfrak{R}_F)_{\text{conv}})_{\text{cl}}$. The final relation (6) is now obvious by Theorem 8, (iii).

DEFINITION 3. *If two almost periodic functions G, F are in the relation* (6), *we shall write*

$$G \dashv F.$$

Remark. It follows from Theorem 9 that the relation $G \dashv F$ is equivalent to $G \epsilon ((\mathfrak{R}_F)_{\text{conv}})_{\text{cl}}$.

THEOREM 10. *$F \dashv G$ and $G \dashv H$ imply $F \dashv H$.*

This follows immediately from Theorem 9.

THEOREM 11. *If $F \epsilon Ap$, and $U \epsilon \mathfrak{U}$, there is a $G \dashv F$, and a number $\mu \geqq 0$, such that for $H \dashv G$ and $a \epsilon \mathfrak{G}$* (cf. [4], Definition 13),

$$\|H(a)\|_U^+ = \mu.$$

We consider, for $K \epsilon ((\mathfrak{R}_F)_{\text{conv}})_{\text{cl}}$, the numerical function

$$\|K\|_{U'}^+ = \underset{x \epsilon \mathfrak{G}}{\text{l.u.b.}}\, \|K(x)\|_U^+. \tag{7}$$

It is a continuous function; hence it assumes, on every compact separable set, a maximum and a minimum; and it is easily seen that, for $K_1 \dashv K_2$, $\|K_1\|_{U'}^+ \leqq \|K_2\|_{U'}^+$. Therefore, if G is an element for which (7) attains its minimum, and if μ denotes the minimum value, we have

$$\|K\|_{U'}^+ = \|H\|_{U'}^+ = \mu, \qquad \text{for } K \dashv H \dashv G.$$

Thus, for $a_1, \cdots, a_n \epsilon \mathfrak{G}; \alpha_1, \cdots, \alpha_n \geqq 0; \alpha_1 + \cdots + \alpha_n = 1$,

$$\text{l.u.b.}_{x\epsilon\mathfrak{G}} \|\alpha_1 H(a_1 x) + \cdots + \alpha_n H(a_n x)\|_U^+ = \text{l.u.b.}_{x\epsilon\mathfrak{G}} \|H(x)\|_U^+ = \mu.$$

Connecting this with

$$\mu \geqq \text{l.u.b.}_{x\epsilon\mathfrak{G}} \|H(x)\|_U^+ \geqq \text{l.u.b.}_{x\epsilon\mathfrak{G}} \sum_{\nu=1}^{n} \alpha_\nu \|H(a_\nu x)\|_U^+,$$

we find for the numerical function $\gamma(x) = \|H(x)\|_U^+$ the property

$$\text{(8)} \qquad \text{l.u.b.}_{x\epsilon\mathfrak{G}} \gamma(x) = \text{l.u.b.}_{x\epsilon\mathfrak{G}} (\alpha_1\gamma(a_1 x) + \cdots + \alpha_n \gamma(a_n x)).$$

By Theorem 5, $\gamma(x)$ is almost periodic; and (8) proves that

$$M_x \gamma(x) = \text{l.u.b.}_{x\epsilon\mathfrak{G}} \gamma(x).$$

(Cf. the definition of M_x in [1], Definition 4.) Hence $\gamma(x)$ is constant (cf. [1], Theorem 7, (4), putting $f(x) = [\text{l.u.b.}_{x\epsilon\mathfrak{G}}\gamma(x)] - \gamma(x)$), and the proof of our theorem is completed.

Corollary. *If $F\epsilon Ap$, $U\epsilon\mathfrak{U}$, $f\epsilon L$, there is a $G \dashv F$, and a number $\mu \geqq 0$, such that for $H \dashv G$, $a\epsilon\mathfrak{G}$,*

$$\text{(9)} \qquad \|H(a) - f\|_U^+ = \mu.$$

Apply Theorem 11 to $F_1(x) = F(x) - f$, and denote the resulting $G_1(x)$ by $G(x) - f$.

Theorem 12. *If $F\epsilon Ap$, there is an $H \dashv F$ which is constant.*

We denote by S the range of $F(x)$, and by T the set $(S_{\text{conv}})_{\text{cl}}$. S is totally bounded (compare the proof of Theorem 6). Hence by [4], Theorems 11, 14, T is totally bounded too. If f (ϵL) is a value of a function $G \dashv F$, then f is a condensation point of elements of the form

$$\alpha_1 F(a_1 x) + \cdots + \alpha_n F(a_n x), \qquad \alpha_1, \cdots, \alpha_n \geqq 0; \alpha_1 + \cdots + \alpha_n = 1;$$

but an element of this form is contained in S_{conv}; thus $f\epsilon T$. Therefore there exists a compact separable set $T \subset L$ which contains the ranges of all $G \dashv F$.

Let $f_1, f_2, \cdots$ be a dense sequence in T, $W_1, W_2, W_3, \cdots$ a complete set of open neighborhoods of zero. Write all pairs $n, p = 1, 2, \cdots$ in a sequence $n_k, p_k, k = 1, 2, \cdots$, and define a sequence of functions $G_0, G_1, G_2, \cdots$ in this manner: $G_0 = F$; G_{k+1} is the G of the corollary to Theorem 11 if applied to $F = G_k$, $U = U_k = W_{p_k}$, $f = f_{n_k}$. Thus

$$F \vdash G_0 \vdash G_1 \vdash G_2 \vdash \cdots,$$

and for all $H \dashv G_k$,

$$\|H(x) - f_{n_k}\|^+_{U_k} = \mu_k, \tag{10}$$

where μ_k is a number depending on k.

The sets $((\mathfrak{R}_{G_k})_{\text{conv}})_{\text{cl}}$ are monotonely decreasing as $k \to \infty$, all closed and non-empty, and all subsets of the compact separable set $((\mathfrak{R}_F)_{\text{conv}})_{\text{cl}}$; thus they have a common element H. Relation (10) means that for any p and any $x, y \in \mathfrak{G}$, the relation

$$\|H(x) - f\|^+_{W_p} = \|H(y) - f\|^+_{W_p} \tag{11}$$

holds for all $f = f_n$. The f_n being dense, and the pseudo-metric being continuous, the relation (11) holds for all $f \in T$. Putting $f = H(y)$ we obtain $\|H(x) - H(y)\|^+_{W_p} = 0$; hence $H(x) - H(y) \in W_p$ $(p = 1, 2, \cdots)$ (cf. [4], Definition 13); hence $H(x) = H(y)$.

COROLLARY. *If $F \in Ap$, there is a $\phi \in L$, and a sequence of systems* $(n = 1, 2, 3, \cdots)$

$$\alpha_{n,1}, \cdots, \alpha_{n,m_n} \geqq 0, \quad \alpha_{n,1} + \cdots + \alpha_{n,m_n} = 1, \quad a_{n,1}, \cdots, a_{n,m_n} \in \mathfrak{G} \tag{12}$$

such that for every $U \in \mathfrak{U}$ there exists an $n_1 = n_1(U)$, for which $n \geqq n_1$ implies

$$\alpha_{n,1}F(a_{n,1}x) + \cdots + \alpha_{n,m_n}F(a_{n,m_n}x) - \phi \in U.$$

We denote the constant value of the function H of Theorem 12 by ϕ. Since $(\mathfrak{R}_F)_{\text{conv}}$ is separable, and H is an element of its closure, H is the limit of a sequence of (equal or different) elements of $(\mathfrak{R}_F)_{\text{conv}}$. Hence the corollary.

THEOREM 13. *If $F \in Ap$, there is a $\psi \in L$, and a sequence of systems* (12), *such that for every $U \in \mathfrak{U}$ there exists an $n_1 = n_1(U)$, for which $n \geqq n_1$ implies*

$$\alpha_{n,1}F(xa_{n,1}y) + \cdots + \alpha_{n,m_n}F(xa_{n,m_n}y) - \psi \in U.$$

Apply the corollary to Theorem 12 to the function $F(z^{-1}y)$ of $L^{\mathfrak{G}^2}$ (cf. Definition 2), and then replace z^{-1} by x.

DEFINITION 4. *If $F \in Ap$, every $\psi \in L$ which has the property described in Theorem 13 is a mean of F.*

THEOREM 14. *If $F, G \in Ap$, and ψ, χ are respective means of F, G, then $\psi \pm \chi$ is a mean of $F \pm G$.*

Given $U \in \mathfrak{U}$, choose $V \in \mathfrak{U}$, with $2V - 2V \in U$. Suppose that $\alpha_\mu \geqq 0$, $\alpha_1 + \cdots + \alpha_m = 1$, $a_\mu \in \mathfrak{G}$, $\beta_\nu \geqq 0$, $\beta_1 + \cdots + \beta_n = 1$, $b_\nu \in \mathfrak{G}$, and that

$$\alpha_1 F(xa_1y) + \cdots + \alpha_m F(xa_my) - \psi \in V, \tag{13}$$

$$\beta_1 F(xb_1y) + \cdots + \beta_n F(xb_ny) - \chi \in V. \tag{14}$$

In (13) we replace y by $b_\nu y$, multiply the equation by β_ν, and sum over ν. We obtain, using [4], Theorem 12,

$$\sum_{\mu=1}^{m}\sum_{\nu=1}^{n}\alpha_\mu\beta_\nu F(xa_\mu b_\nu y) - \psi\epsilon\sum_{\nu=1}^{n}\beta_\nu V \subset 2V.$$

Similarly, if in (14) we replace x by xa_μ, multiply by α_μ and sum over μ, we get

$$\sum_{\mu=1}^{m}\sum_{\nu=1}^{n}\alpha_\mu\beta_\nu G(xa_\mu b_\nu y) - \chi\epsilon\sum_{\mu=1}^{m}\alpha_\mu V \subset 2V.$$

Hence

$$\sum_{\mu=1}^{m}\sum_{\nu=1}^{n}\alpha_\mu\beta_\nu(F(xa_\mu b_\nu y) - G(xa_\mu b_\nu y)) - (\psi - \chi)\epsilon 2V - 2V \subset U.$$

As

$$\alpha_\mu\beta_\nu \geqq 0,\ \sum_{\mu=1}^{m}\sum_{\nu=1}^{n}\alpha_\mu\beta_\nu \geqq 1,\ a_\mu b_\nu \epsilon \mathfrak{G},$$

and U was arbitrary, we conclude that $\psi-\chi$ is a mean of $F-G$. A similar argument proves that $\psi+\chi$ is a mean of $F+G$.

COROLLARY. *If $F_1, F_2, \cdots, F_N\epsilon Ap$ $(N=1, 2, 3, \cdots)$, with the means $\psi_1, \psi_2, \cdots, \psi_N$ respectively, and if $U\epsilon\mathfrak{U}$, then there exist numbers $\alpha_1, \cdots, \alpha_n \geqq 0$, $\alpha_1+\cdots+\alpha_n=1$, and elements $a_1, \cdots, a_n\epsilon\mathfrak{G}$, such that simultaneously for $\nu=1, 2, \cdots, N$,*

$$\alpha_1F_\nu(xa_1y) + \alpha_2F_\nu(xa_2y) + \cdots + \alpha_nF_\nu(xa_ny) - \psi_\nu \subset U.$$

The case $N=2$ was treated in the proof of Theorem 14, and the same argument can be used to extend our statement from N to $N+1$.

THEOREM 15. *Every almost periodic function has one and only one mean.*

If ψ and χ are both means of F, $\psi-\chi$ is a mean of $F-F=0$. But every mean of 0 is 0. Hence the uniqueness of the mean.

DEFINITION 5. *If $F\epsilon Ap$, its (unique) mean will be denoted by MF or $M_xF(x)$.*

THEOREM 16. *The mean has the following properties:*

1. *If $F(x)=\phi$ (constant ϵL), then $MF=\phi$.*
2. *If α is a number, $M(\alpha F)=\alpha MF$.*
3. *$M(F\pm G)=MF\pm MG$.*
4. *$M_xF(ax)=M_xF(x)$.*
5. *$M_xF(xa)=M_xF(x)$.*
6. *$M_xF(x^{-1})=M_xF(x)$.*
7. *If $U\epsilon\mathfrak{U}$, $\|MF\|_U^+ \leqq M_x(\|F(x)\|_U^+) \leqq \|F\|_U^+$.*
8. *If $U\epsilon\mathfrak{U}$, and $F-G\epsilon U$, then $MF-MG\epsilon 2U$.*

1 and 2 are obvious. 3 follows from Theorem 14. 4, 5, 6 are easily deduced from Definition 5. The first half of 7 follows from the relation

$$\|\alpha_1 F(xa_1y) + \cdots + \alpha_n F(xa_ny)\|_U^+ \leqq \alpha_1\|F(xa_1y)\|_U^+ + \cdots + \alpha_n\|F(xa_ny)\|_U^+$$

and the continuity of the pseudo-metric; the second half is a proved theorem on real almost periodic functions (cf. [1], Theorem 7). With regard to 8, from $\|F-G\|_{U'}^+ \leqq 1$, it follows that $\|MF-MG\|_U^+ \leqq 1$, and thus $MF-MG\epsilon 2U$.

THEOREM 17. *The properties* 1, 2, 3, 4 (*or* 5), 8, *of Theorem* 16, *determine our mean uniquely.*

Replacing ab by ba in $\mathfrak{G}$ shows that it is sufficient to consider the properties 1, 2, 3, 4, 8. Let NF be a notion defined in Ap satisfying these properties. If $V+V\subset U\epsilon\mathfrak{U}$, $\alpha_1+\cdots+\alpha_n=1$, and

$$\alpha_1 F(a_1x) + \cdots + \alpha_n F(a_nx) - MF \subset V,$$

then

$$NF - MF = N(\alpha_1 F(a_1x) + \cdots + \alpha_n F(a_nx) - MF) \subset 2V \subset U.$$

U being arbitrary, $NF=MF$.

Remark. The properties 3, 7 of Theorem 16 imply the property 8. Hence we have that the properties 1, 2, 3, 4 (or 5), 7, of Theorem 16 determine our mean uniquely.

THEOREM 18. *If* $x\epsilon\mathfrak{G}$, $y\epsilon\mathfrak{G}'$, *and* $F(x, y)$ *is almost periodic in* x, y, *then* $F(x, y)$ *is almost periodic in* x *for fixed* y, *and almost periodic in* y *for fixed* x, $G(x)=M_yF(x, y)$ *is almost periodic in* x, $H(y)=M_xF(x, y)$ *is almost periodic in* y, *and*

$$\begin{aligned} MF &= M_xG(x) = M_x(M_yF(x, y)), \\ MF &= M_yH(y) = M_y(M_xF(x, y)). \end{aligned} \tag{15}$$

If the sets $F(ax, by)$, $F(xa, yb)$ are totally bounded for $a\epsilon\mathfrak{G}$, $b\epsilon\mathfrak{G}'$, then the sets $F(ax, y)$, $F(xa, y)$, which are parts of them, are totally bounded for $a\epsilon\mathfrak{G}$. Hence $F(x, y)$ is almost periodic in x for fixed y. Similarly, if we interchange x and y, $F(x, y)$ is almost periodic in y for fixed x.

Given $U\epsilon\mathfrak{U}$, choose $V\epsilon\mathfrak{U}$, with $V+V\subset U$, and elements $a_1, \cdots, a_m, b_1, \cdots, b_n$, such that

$$F(ax, y)\epsilon\mathfrak{S}(F(a_1x, y) + V', \cdots, F(a_mx, y) + V'),$$
$$F(xb, y)\epsilon\mathfrak{S}(F(xb_1, y) + V', \cdots, F(xb_n, y) + V').$$

By Theorem 16, 8,

$$G(ax) = M_yF(ax, y)\epsilon\mathfrak{S}(M_yF(a_1x, y) + U', \cdots, M_yF(a_mx, y) + U'),$$

and similarly

$$G(xb)\epsilon\mathfrak{S}(M_yF(xb_1, y) + U', \cdots, M_yF(xb_n, y) + U').$$

Hence $G(x)$ is almost periodic, and similarly $H(y)$. In order to prove the first half of (15) we have only to show that the "mean" $NF=M_x(M_yF)$ satisfies the conditions 1, 2, 3, 4, 7, of Theorem 16. But this is easily verified; for instance, in the case of the first half of condition 7, we have

$$\|NF\|_U^+=\|M_x(M_yF(x, y))\|_U^+ \leqq M_x(\|M_yF(x, y)\|_U^+) \leqq M_xM_y(\|F\|_{U'}^+) = N(\|F\|_{U'}^+)$$

(cf. [1], Theorem 10). Similarly for the second half of (15).

Part II. Fourier expansions

Definition 6. *If $\phi(x)$ is a numerical almost periodic function, and $F\epsilon Ap$, $\phi\times F$ is the function*

$$M_y(\phi(xy^{-1})F(y)) = M_y(\phi(xy)F(y^{-1}))$$

which again belongs to Ap.

Theorem 19. *$\phi\times F$ is linear in ϕ and in F, and associative: $(\phi\times\psi)\times F=\phi\times(\psi\times F)$.*

The first follows from Theorem 16, 2, 3. The second follows from the following formal calculations, each step of which is justified by one of the foregoing theorems:

$$\begin{aligned}(\phi \times \psi) \times F &= M_y(M_z(\phi(xy^{-1}z^{-1})\psi(z))F(y)) = M_yM_z(\phi(xy^{-1}z^{-1})\psi(z)F(y))\\ &= M_{(y,z)}(\phi(xy^{-1}z^{-1})\psi(z)F(y)).\\ \phi \times (\psi \times F) &= M_y(\phi(xy^{-1})M_z(\psi(yz^{-1})F(z))) = M_yM_z(\phi(xy^{-1})\psi(yz^{-1})F(z))\\ &= M_zM_y(\phi(xy^{-1})\psi(yz^{-1})F(z)) = M_z(M_y(\phi(xy^{-1})\psi(yz^{-1}))F(z)),\end{aligned}$$

and substituting yz for y,

$$\begin{aligned}&= M_z(M_y(\phi(xz^{-1}y^{-1})\psi(y))F(z)) = M_zM_y(\phi(xz^{-1}y^{-1})\psi(y)F(z))\\ &= M_{(y,z)}(\phi(xy^{-1}z^{-1})\psi(z)F(y)).\end{aligned}$$

Theorem 20. *If $F(x, y)$ is an almost periodic function of $L^{\mathfrak{G}\times\mathfrak{G}'}$, and $U\epsilon\mathfrak{U}$, there exist numbers $\alpha_1, \cdots, \alpha_n \geqq 0$, $\alpha_1+\cdots+\alpha_n=1$, and elements $a_\nu\epsilon\mathfrak{G}$, such that, for all $x\epsilon\mathfrak{G}$, $y\epsilon\mathfrak{G}'$,*

$$\sum_{\nu=1}^{n} \alpha_\nu F(a_\nu x, y) - M_xF(x, y)\epsilon U. \tag{16}$$

For fixed y, $F(x, y)$ is almost periodic in x. Hence, by the corollary to Theorem 14, if any finite number of fixed values is assigned to the variable y, it is possible to choose quantities α_ν, a_ν which satisfy (16) for all x.

Now, choose $V\epsilon\mathfrak{U}$, with $2V+V-2V\subset U$, and then elements $b_1, \cdots, b_m$, such that $F(x, yb)\epsilon\mathfrak{S}(F(x, yb_1)+V', \cdots, F(x, yb_m)+V')$. Putting $y=1$, we obtain

$$F(x, b)\epsilon\mathfrak{S}(F(x, b_1) + V', \cdots, F(x, b_m) + V').$$

We determine quantities α_ν, a_ν such that

$$\sum_{\nu=1}^{n} \alpha_\nu F(a_\nu x, b_\mu) - M_x F(x, b_\mu)\epsilon V \tag{17}$$

for $\mu=1, \cdots, m$. To each b there corresponds a b_μ such that

$$F(x, b) - F(x, b_\mu)\epsilon V. \tag{18}$$

Hence by Theorem 16, 8,

$$M_x F(x, b) - M_x F(x, b_\mu)\epsilon 2V. \tag{19}$$

On the other hand, it follows from (18) that

$$\sum_{\nu=1}^{n} \alpha_\nu F(a_\nu x,b) - \sum_{\nu=1}^{n} \alpha_\nu F(a_\nu x, b_\mu)\epsilon 2V. \tag{20}$$

Combining (17), (19), (20), we obtain, for any $b\epsilon\mathfrak{G}'$,

$$\sum_{\nu=1}^{n} \alpha_\nu F(a_\nu x, b) - M_x F(x, b)\epsilon 2V + V - 2V\subset U,$$

and this proves the theorem.

DEFINITION 7. *A weight function is a real almost periodic function $\phi(x)$ with the properties: $\phi(x)\geqq 0$, $M\phi=1$. A special weight function is a weight function which is a finite linear aggregate of representation coefficients.* (Cf. [1], Chapter III.)

THEOREM 21. *If $F\epsilon Ap$, and $\phi(x)$ is a weight function, then $\phi\times F\dashv F$.*

If $U\epsilon\mathfrak{U}$, choose $V\epsilon\mathfrak{U}$, with $V+V\subset U$. Then $\phi\times F=M_y(\phi(y)F(y^{-1}x))$. By Theorem 20 there are numbers $\alpha_1, \cdots, \alpha_n\geqq 0$, $\alpha_1+\cdots+\alpha_n=1$, and elements $a_1, \cdots, a_n\epsilon\mathfrak{G}$, such that

$$\phi \times F - \sum_{\nu=1}^{n} \alpha_\nu\phi(a_\nu y)F(y^{-1}a_\nu^{-1}x)\epsilon V. \tag{21}$$

The function

$$\psi(y) = \sum_{\nu=1}^{n} \alpha_\nu\phi(a_\nu y)$$

is a weight function. Hence there exist elements y_1, y_2, such that $\psi(y_1) \geqq 1$, $\psi(y_2) \leqq 1$. Therefore we may find a γ, $0 \leqq \gamma \leqq 1$, with $\gamma\psi(y_1)+(1-\gamma)\psi(y_2)=1$. From (20) we easily deduce that

$$\phi \times F \subset \gamma \sum_{\nu=1}^{n} \alpha_\nu \phi(a_\nu y_1) F(y_1^{-1} a_\nu^{-1} x) + (1-\gamma) \sum_{\nu=1}^{n} \alpha_\nu \phi(a_\nu y_2) F(y_2^{-1} a_\nu^{-1} x) + U'.$$

It is easily seen that the function in x on the right side $\epsilon(\mathfrak{R}_F)_{\text{conv}}$. Hence $\phi \times F \epsilon (\mathfrak{R}_F)_{\text{conv}} + U'$, for any $U' \epsilon \mathfrak{U}'$. Thus ([4], Theorem 5) $\phi \times F \subset ((\mathfrak{R}_F)_{\text{conv}})_{\text{cl}}$.

Lemma 1. *If $\phi(x, y)$ is a real almost periodic function in x, y, then*

$$\psi(x) = \underset{y \epsilon \mathfrak{G}'}{\text{l.u.b.}}\ \phi(x, y)$$

is almost periodic in x.

If $|\phi(ax, by) - \phi(a_\nu x, b_\nu y)| \leqq \epsilon$, then also

$$\left| \underset{y \epsilon \mathfrak{G}'}{\text{l.u.b.}}\ \phi(ax, by) - \underset{y \epsilon \mathfrak{G}'}{\text{l.u.b.}}\ \phi(a_\nu x, b_\nu y) \right| \leqq \epsilon,$$

that is,

$$\left| \underset{y \epsilon \mathfrak{G}'}{\text{l.u.b.}}\ \phi(ax, y) - \underset{y \epsilon \mathfrak{G}'}{\text{l.u.b.}}\ \phi(a_\nu x, y) \right| \leqq \epsilon.$$

Hence, if $\mathfrak{R}_\phi$ is totally bounded, then so is $\mathfrak{R}_\psi$. Similarly for $\mathfrak{L}_\psi$.

Lemma 2. *If $\psi(x)$ is a weight function, and $\epsilon > 0$, there is a special weight function $\chi(x)$ such that*

$$|\psi(x) - \chi(x)| \leqq \epsilon, \qquad x \epsilon \mathfrak{G}.$$

By the approximation theorem ([1], Theorem 30), for each $\delta > 0$, there is a finite linear aggregate of representation coefficients $\chi_1(x)$ such that

$$|\psi(x) - \chi_1(x)| \leqq \delta. \tag{22}$$

We may assume $\chi_1(x)$ real, otherwise we replace it by $(\chi_1(x)+\overline{\chi_1(x)})/2$. After that, we may assume it $\geqq 0$, otherwise $\chi_1(x)+\delta$ satisfies (22) with 2δ instead of δ. Since $M\psi = 1$, (22) implies $1-\delta \leqq M\chi_1 \leqq 1+\delta$; hence, for $\chi = \chi_1/(M\chi_1)$, we obtain

$$|\psi(x) - \chi(x)| \leqq |\psi(x) - \chi_1(x)| + \left| \left(1 - \frac{1}{M\chi_1}\right)\chi_1 \right|$$
$$\leqq \delta + \frac{\delta}{1-\delta}\left[\underset{x \epsilon \mathfrak{G}}{\text{l.u.b.}}\ \psi(x) + \delta\right],$$

and this is $\leqq \epsilon$, if δ is small enough.

THEOREM 22. *If $F \epsilon Ap$, and $U \epsilon \mathfrak{U}$, there exists a special weight function χ such that $\chi \times F - F \epsilon U'$.*

$F(x^{-1}y) - F(y)$ is almost periodic in x, y, hence $\|F(x^{-1}y) - F(y)\|_U^+$ is also almost periodic in x, y (Theorem 5). By Lemma 1,

$$t(x) = \operatorname*{l.u.b.}_{y \epsilon \mathfrak{G}} \|F(x^{-1}y) - F(y)\|_U^+$$

is almost periodic in x.

Given $\epsilon > 0$, we form with the function

$$\omega_\epsilon(u) = \begin{cases} 1 - \dfrac{|u|}{\epsilon} & \text{for } |u| \leqq \epsilon, \\ 0 & \text{for } |u| \geqq \epsilon \end{cases}$$

the almost periodic (Theorem 5) function $\psi_\epsilon(x) = \omega_\epsilon(t(x))$. We have (1) $\psi_\epsilon(x) \geqq 0$, (2) $\psi_\epsilon(1) = 1$, since $t(0) = 0$, (3) if $\psi_\epsilon(x) > 0$, then $\|F(x^{-1}y) - F(y)\|_U^+ \leqq \epsilon$ for all $y \epsilon \mathfrak{G}$, (4) $m_\epsilon = M_x \psi_\epsilon(x) > 0$ by [1], Theorem 7, 4. Therefore

$$\|\psi_\epsilon(z)(F(z^{-1}x) - F(x))\|_U^+ \leqq \psi_\epsilon(z)\|F(z^{-1}x) - F(x)\|_U^+ \leqq \psi_\epsilon(z)t(z),$$

and this is equal to 0 if $\psi_\epsilon(z) = 0$, and $\leqq \epsilon\psi_\epsilon(z)$ if $\psi_\epsilon(z) \neq 0$. Thus

$$\|\psi_\epsilon \times F - m_\epsilon F\|_U^+ \leqq \epsilon m_\epsilon,$$

and for the weight function $\psi(x) = \psi_\epsilon(x)/m_\epsilon$ we have

$$\|\psi \times F - F\|_U^+ \leqq \epsilon.$$

For the special weight function $\chi(x)$ from Lemma 2 we obtain

$$\|\chi \times F - F\|_U^+ \leqq \epsilon + \|(\chi - \psi) \times F\|_U^+ \leqq \epsilon + M_x\|(\psi(xy^{-1}) - \chi(xy^{-1}))F(y)\|_U^+$$

$$\leqq \epsilon + \epsilon M_x(\max \| \pm F(y)\|_U^+) \leqq \epsilon + \epsilon C,$$

where C is a constant independent of ϵ. Choosing $\epsilon < (1+C)^{-1}$, we obtain the theorem.

THEOREM 23. *If $F \epsilon Ap$, there exists a sequence of special weight functions $\phi_n(x)$, $n = 1, 2, 3, \cdots$, such that F is the limit of the sequence $\phi_n \times F$, $n = 1, 2, 3, \cdots$.*

Consider the functions $\phi \times F$, for all special weight functions ϕ, and denote their set by S. By Theorem 22, $F \subset S_{cl}$; but S_{cl}, being $\subset ((\mathfrak{R}_F)_{conv})_{cl}$ (Theorem

21), is separable; hence, F is the limit of a sequence of (equal or different) elements ϵS.

DEFINITION 8. *If* $D(x) = \{D_{\rho\sigma}(x)\}_{\rho,\sigma=1,\cdots,s}$ *is an irreducible normal representation of* $\mathfrak{G}$ (cf. [1], Chapter III, Definitions 9, 10) *we form, for* $F \epsilon Ap$,

$$f_{\rho\sigma}(D) = M_x(D_{\sigma\rho}(x^{-1})F(x)) = M(\overline{D_{\rho\sigma}}F),$$

and the matrix

$$f^D = \{f_{\rho\sigma}(D)\}_{\rho,\sigma=1,\cdots,s}.$$

$f_{\rho\sigma}(D)$ *is the* (D, ρ, σ)*-expansion-coefficient of* F, f^D *its D-expansion-matrix.*

THEOREM 24. *The expansion coefficients and matrices have the following properties:*

1. $M(\overline{D_{\rho\sigma}}(\alpha F)) = \alpha M(\overline{D_{\rho\sigma}}F)$, $\alpha \epsilon L$.
2. $M(\overline{D_{\rho\sigma}}(F+G)) = M(\overline{D_{\rho\sigma}}F) + M(\overline{D_{\rho\sigma}}G)$.
3. *If* F *is the limit of a sequence* F_N, $N=1, 2, \cdots$, *then, for each* (D, ρ, σ), $M(\overline{D_{\rho\sigma}}F)$ *is the limit of the sequence* $M(\overline{D_{\rho\sigma}}F_N)$, $N=1, 2, \cdots$.
4. *If* $G=\phi\times F$, *and if* $g_{\rho\sigma}(D)$, $\phi_{\rho\sigma}(D)$, $f_{\rho\sigma}(D)$ *are expansion coefficients of* G, ϕ, F *respectively, then*

$$g^D = \phi^D f^D,$$

that is to say,

$$g_{\rho\sigma}(D) = \sum_{\tau=1}^{s} \phi_{\rho\tau}(D) f_{\tau\sigma}(D).$$

In particular, the matrix g^D *vanishes if* f^D *or* ϕ^D *vanishes (or both).*

5. *If* D^N, $N=1, 2, 3, \cdots$, *is a (finite or countable) sequence of irreducible inequivalent normal representations, if the* s^N *are their respective degrees, if* $h_{\rho\sigma}$ *are elements of* L, *and if*

$$F = \sum_{N=1}^{\infty}\left(s^N \sum_{\rho,\sigma=1}^{s^N} \overset{N}{h}_{\rho\sigma} \overset{N}{D}_{\rho\sigma}\right) \tag{23}$$

[*meaning that* $F\epsilon Ap$ *is the limit,* $m\to\infty$, *of the sequence of elements*

$$F_m = \sum_{N=1}^{m}\left(s^N \sum_{\rho,\sigma=1}^{s^N} \overset{N}{h}_{\rho\sigma} \overset{N}{D}_{\rho\sigma}\right)$$

from Ap], *then*

$$M(\overline{D_{\rho\sigma}}F) = \begin{cases} 0 \text{ if } D \neq D^N \\ \overset{N}{h}_{\rho\sigma} \text{ if } D = D^N \end{cases} \qquad (N = 1, 2, 3, \cdots).$$

1 and 2 are obvious; 3 is easily deducible from the fact that $G(x, y)\epsilon U$ implies $M_yG(x, y)\epsilon 2U$ (Theorem 16, 8); and 4 follows from the relation

$$\begin{aligned} \overset{D}{g}_{\rho\sigma} &= M_x(D_{\sigma\rho}(x^{-1})M_y(\phi(xy^{-1})F(y))) = M_y(F(y)M_x(D_{\sigma\rho}(x^{-1})\phi(xy^{-1}))) \\ &= M_y(F(y)M_z(D_{\sigma\rho}(y^{-1}z^{-1})\phi(z))) \\ &= \sum_{\tau=1}^{n} M_y(F(y)D_{\sigma\tau}(y^{-1}))\cdot M_z(D_{\tau\rho}(z^{-1})\phi(z)). \end{aligned}$$

As regards 5, if the number of the representations D^N is finite, it follows by direct computation of $M_x(\overline{D_{\rho\sigma}}(x)F(x))$, applying 1 and 2 and [1], Theorem 21, and in the general case by using this and 3.

THEOREM 25. *If $F\epsilon Ap$, it has only countably many expansion matrices $\neq 0$.*

By Theorem 24, 4 and 3, this is true if F has the form $\phi\times G$, with ϕ a special weight function, or if F is the limit of such functions. Now apply Theorem 23.

DEFINITION 9. *If $F\epsilon Ap$; if $D^N, N=1, 2, \cdots$, is the sequence of irreducible normal representations (in any fixed order) for which its expansion matrices $\neq 0$; if the s^N are their respective degrees; and if $f_{N,\rho\sigma}=M_x(D^N_{\rho\sigma}(x^{-1})F(x))$; then we call the formal series*

$$\sum_{N=1}^{\infty} s^N \sum_{\rho,\sigma=1}^{s^N} f_{N,\rho\sigma} D^N_{\rho\sigma} \tag{24}$$

the Fourier expansion of F.

We call a sequence F_m, $m=1, 2, \cdots$, formally convergent (to F), if the sequence of the Fourier expansions of the functions F_m is formally convergent (to the Fourier expansion of F), i.e., if for any (D, ρ, σ), the sequence $M(\overline{D_{\rho\sigma}}F_m)$ has a limit (namely $M(\overline{D_{\rho\sigma}}F)$).

Remark. Theorem 24 states properties of the Fourier expansion. 1 and 2 state additivity. 3 states that the Fourier series of a limit is the formal limit of the Fourier series. 4 gives an important rule for the computation of the Fourier series of a convolution of an almost periodic function with a numerical almost periodic function. Finally, 5 states that the sum of a uniformly convergent series of the form (23) is its own Fourier expansion.

THEOREM 26. (Uniqueness Theorem.) *Almost periodic functions which have the same Fourier expansion are equal.*

If G and H have the same expansion, then the expansion of $F=G-H$ vanishes identically. From

$$D_{\rho\sigma} \times F = M_y(D_{\rho\sigma}(xy^{-1})F(y)) = \sum_{\tau=1}^{s} D_{\rho\tau}(x)M_y(D_{\tau\sigma}(y^{-1})F(y)) \tag{25}$$

it follows that $D_{\rho\sigma}\times F$ vanishes for any (D, ρ, σ). Hence $\phi\times F=0$ for any special weight function. But F is the limit of such functions (Theorem 23). Therefore $F=0$.

Part III. Theorems on approximation and summation

Theorem 27. (Approximation Theorem.) *If F is almost periodic, it is the limit of a sequence of finite linear aggregates of the form $\sum fD_{\rho\sigma}$ with $f\epsilon L$. More precisely, if D^N, $N=1, 2, 3, \cdots$, are the representations occurring in the Fourier expansion of F, if the s^N are their respective degrees, there exist elements $f^m_{N,\rho\sigma}$ of L $(m=1, 2, 3, \cdots)$ such that for each m only a finite number of them is $\neq 0$, and that F is the limit, for $m\to\infty$, of the finite aggregates*

$$F_m = \sum_{N=1}^{\infty}\left(s^N \sum_{\rho,\sigma=1}^{s^N} f^m_{N,\rho\sigma}D^N_{\rho\sigma}\right).$$

F is the limit of a sequence $F_m=\phi_m\times F$, each ϕ_m being a special weight function. Using (25) we find that this sequence has the property stated in the theorem.

Theorem 28. *Let the sequence F_m, $m=1, 2, 3, \cdots$, be part of a totally bounded or compact set of Ap. In order that the sequence have a limit it is sufficient that it be formally convergent.*

As the closure of a totally bounded set is compact (cf. [4], Theorem 11 and Definition 10), it is sufficient to consider the second case. Owing to the compactness, the sequence has at least one condensation point, and, using the compactness again, we have to show that it has no more than one condensation point. Otherwise there would exist two subsequences F_{p_m}, F_{q_m} of F_m, having two different limits, G and H respectively. The Fourier expansion of G is the formal limit of the Fourier expansions of the sequence F_{p_m}, and therefore, the sequence F_m being formally convergent, of the sequence F_m. Similarly, the Fourier expansion of H is the formal limit of the expansions of the sequence F_m. Thus G and H have the same Fourier expansion, and by Theorem 26, $G=H$, which contradicts our hypothesis.

Definition 10. *The function $F\epsilon Ap$ will be called a class function if for all x and y, $F(yxy^{-1})=F(x)$, or, which is equivalent, if for all x and y, $F(yx)=F(xy)$.*

A formal series (24) *will be called a class series if*

$$f_{N,\rho\sigma} = f_N\delta_{\rho\sigma}. \tag{26}$$

THEOREM 29. *If $F(x)\epsilon Ap$, then $F_0(x)=M_yF(yxy^{-1})$ is a class function, and if* (24) *is the Fourier expansion of $F(x)$, then the Fourier expansion of $F_0(x)$ is*

$$\sum_{N=1}^{\infty} s^N f_N \sum_{\rho=1}^{s^N} D_{\rho\rho}^N,$$

where

$$f_N = \frac{1}{s^N}\sum_{\rho=1}^{s^N} f_{N,\rho\rho}.$$

In order that a function F of Ap be a class function, it is necessary and sufficient that its Fourier expansion be a class series.

We have

$$F_0(zxz^{-1}) = M_y F(yzxz^{-1}y^{-1}) = M_y F(yxy^{-1}) = F_0(x);$$

hence $F_0(x)$ is a class function. For a fixed representation $\{D_{\rho\sigma}\}$ we have

$$\begin{aligned} M_x(D_{\rho\sigma}(x^{-1})F_0(x)) &= M_xM_y(D_{\rho\sigma}(x^{-1})F(yxy^{-1})) = M_xM_y(D_{\rho\sigma}(y^{-1}x^{-1}y)F(x)) \\ &= M_y \sum_{p,q=1}^{s} D_{\rho p}(y^{-1})D_{q\sigma}(y)\cdot M_x(D_{pq}(x^{-1})F(x)) \\ &= \frac{\delta_{\rho\sigma}}{s}\sum_{p=1}^{s} M_x(D_{pp}(x^{-1})F(x)) \end{aligned}$$

and this proves the statement about the Fourier series of $F_0(x)$.

If the Fourier expansion of F is a class series, the functions F and F_0 have the same Fourier expansion. By Theorem 26, $F=F_0$, and hence F is a class function, because F_0 is one.

Conversely if $F(x)=F(yxy^{-1})$, then $F(x)=M_yF(yxy^{-1})=F_0(x)$, and by the first part of our theorem, the Fourier expansion of $F_0(x)$ is a class series.

THEOREM 30. (Summation Theorem.) *Let D^N, $N=1, 2, 3, \cdots$, denote a sequence of irreducible normal representations, and the s^N their respective degrees. There exists a sequence of special weight functions ϕ_m, $m=1, 2, 3, \cdots$, with the following properties:*

1. *Each ϕ_m is a class function.*
2. *All Fourier coefficients of ϕ_m are $\geqq 0$, $\leqq 1$.*
3. *Any almost periodic function F whose Fourier expansion contains no other representations than the given ones, is the limit of the sequence $\phi_m \times F$, $m=1, 2, \cdots$.*

In particular, there exists a square array of numbers r_N^m, m, $N=1, 2, \cdots$, with the following properties:

α. *For each m only a finite number of them is $\neq 0$.*
β. $0 \leqq r_N^m \leqq 1$.
γ. *If an almost periodic function F has a Fourier expansion*

$$\sum_{N=1}^{\infty} s^N \sum_{\rho,\sigma=1}^{s^N} f_{N,\rho\sigma} D_{\rho\sigma}^N$$

[*any number of the coefficients $f_{N,\rho\sigma}$ may vanish*], *then F is the limit, $m \to \infty$, of the finite aggregates*

$$F_m = \sum_{N=1}^{\infty} r_N^m\, s^N \sum_{\rho,\sigma=1}^{s^N} f_{N,\rho\sigma} D_{\rho\sigma}^N.$$

We determine numbers ϵ_N, all $\neq 0$, such that the series

$$\sum_{N=1}^{\infty} \epsilon_N s^N \left(\sum_{\tau=1}^{s^N} D_{\tau\tau}^N(x) \right) \tag{27}$$

is uniformly convergent, thus representing a numerical almost periodic function $f(x)$. There exist special weight functions $\chi_m(x)$, $m = 1, 2, \cdots$, such that the sequence of functions $f_m(x) = \chi_m \times f(x)$ is uniformly convergent to $f(x)$. By [1], Theorem 21,

$$f \times D_{\rho\sigma}^N = \epsilon_N D_{\rho\sigma}^N.$$

But $f_m \times D_{\rho\sigma}^N = \chi_m \times (f \times D_{\rho\sigma}^N)$. Considering $\epsilon_N \neq 0$, we conclude that

$$\lim_{m \to \infty} \chi_m \times D_{\rho\sigma}^N = D_{\rho\sigma}^N. \tag{28}$$

In particular,

$$\lim_{m \to \infty} M_x(D_{\rho\sigma}^N(x^{-1})\chi_m(x)) = \delta_{\rho\sigma}.$$

We now consider the class function

$$\psi_m(x) = M_y \chi_m(yxy^{-1}); \tag{29}$$

$\psi_m(x)$ is obviously a weight function, and Theorem 29 shows that it has the form

$$\psi_m(x) = \sum_D s_D\, a_D^m \left(\sum_{\tau=1}^{s_D} D_{\tau\tau} \right) \tag{30}$$

with a finite number of terms, so that it is a special weight function. From (28) it follows that

$$\lim_{m\to\infty} \overset{m}{a}_D = 1 \qquad \text{for } D = D^N.$$

The function

$$\psi_m' = \overline{\psi}_m(x^{-1}) = \sum_D s_D \overline{\overset{m}{a}_D}\left(\sum_{\tau=1}^{s_D} D_{\tau\tau}\right)$$

has all the properties of ψ_m. Finally we introduce the class function $\phi_m = \psi_m \times \psi_m'$, for which we find

$$\phi_m = \sum_D s_D \overset{m}{r}_D\left(\sum_{\tau=1}^{s_D} D_{\tau\tau}\right),$$

where

$$\overset{m}{r}_D = \left|\overset{m}{a}_D\right|^2,$$

with

(31) $$\lim_{m\to\infty} \overset{m}{r}_D = 1 \qquad \text{for } D = D^N.$$

It is easy to find that ϕ_m is again a weight function. Its coefficients $\overset{m}{r}_D$ are $\geqq 0$, and as $\overset{m}{r}_D$, $\phi_m(x)$ are real and $\geqq 0$,

$$\overset{m}{r}_D = M_x(D_{11}(x^{-1})\phi_m(x)) \leqq M_x(\left|D_{11}(x^{-1})\right| \cdot \left|\phi_m(x)\right|) \leqq M_x(\phi_m(x)) = 1.$$

Hence the properties 1 and 2 of our theorem are fulfilled. Using (31) we conclude from Theorem 24, 4, that for any F whose Fourier expansion contains no other representations than the given ones, the sequence $\phi_m \times F$ converges formally to F. By Theorem 21, this sequence is part of the compact set $((\mathfrak{R}_F)_{\mathrm{conv}})_{\mathrm{cl}}$. Thus we may apply Theorem 28, and this proves property 3.

The second half of Theorem 30 is an immediate consequence of the first half.

Definition 11. *A system of irreducible normal representations will be called a module $\mathfrak{M}$ if it contains with every representation its complex conjugate, and the Fourier series of the product of any two occurring representation coefficients contains no other representations than the given ones.* A module will be called countable if it contains only a finite or enumerable number of representations.*

* If D^M, D^N are any two normal representations, then, in the sense in which these symbols are used in the theory of group-representations, the direct product $D^M D^N$ is a finite sum

$$\sum_P c_P^{MN} D^P$$

(the c_P^{MN} are the "composition coefficients"). Hence the product of any two representation coefficients has a *finite* Fourier expansion.

LEMMA 3. *There is always a smallest module containing a given set of representations; if the given set is countable, then so is also the smallest module.*

Consider all finite subsystems of the given set of representation coefficients and their complex conjugates and form all possible power-products with them. The totality of the representations occurring in their Fourier series is the desired module.

DEFINITION 12. *Given any module $\mathfrak{M}$, we shall denote by $\{\mathfrak{M}\}$ the set of Ap consisting of those functions whose Fourier expansion contains no other representations than those occurring in $\mathfrak{M}$.*

THEOREM 31. *Given $\mathfrak{M}$, the set $\{\mathfrak{M}\}$ has the following properties:*

1. *If $F(x)$ is in $\{\mathfrak{M}\}$, every $F(xa)$ is in $\{\mathfrak{M}\}$.*
2. *If $F(x)$ is in $\{\mathfrak{M}\}$, every $F(ax)$ is in $\{\mathfrak{M}\}$.*
3. *If $F(x)$ is in $\{\mathfrak{M}\}$, every $\alpha F(x)$ is in $\{\mathfrak{M}\}$.*
4. *If $F(x)$ and $G(x)$ are in $\{\mathfrak{M}\}$, $F(x)\pm G(x)$ is in $\{\mathfrak{M}\}$.*
5. *If $F_1, F_2, \cdots$ are in $\{\mathfrak{M}\}$, and if F is the limit of F_n, then F is in $\{\mathfrak{M}\}$.*
6. *If $f_1, \cdots, f_n$ are numerical functions of $\{\mathfrak{M}\}$, and $F(u_1, \cdots, u_n)$ is a numerical function which is defined and uniformly continuous for the range of $f_1, \cdots, f_n$, then $F(f_1(x), \cdots, f_n(x))$ is also contained in $\{\mathfrak{M}\}$.*
7. *If at least one of the two functions $\alpha(x), F(x)$ is in $\{\mathfrak{M}\}$, then $\alpha\times F(x)$ is also. ($\alpha(x)$ is numerical.)*

1–5 follow easily from the formal properties of Fourier expansion. 7 is an immediate consequence of Theorem 24, 4. As regards 6, if f and g are numerical finite aggregates of representations from $\{\mathfrak{M}\}$, then $\bar{f}$ as well as fg is also $\epsilon\{\mathfrak{M}\}$ by definition of $\{\mathfrak{M}\}$. Applying the approximation theorem and 5, the same holds for $\bar{f}$ and fg for any numerical functions f, g from $\{\mathfrak{M}\}$. Hence 6 is true if $F(u_1, \cdots, u_n)$ is a polynomial in $u_1, \cdots, u_n$ and their complex conjugates. $F(u_1, \cdots, u_n)$ being uniformly continuous on the range of $f_1, \cdots, f_n$, which is a bounded set since $f_1, \cdots, f_n$ are a.p., $F(f_1(x), \cdots, f_n(x))$ is a uniform limit of polynomials in $f_1(x), \cdots, f_n(x)$ and their complex conjugates. Thus 5 completes the proof of 6.

THEOREM 32. *If $\mathfrak{M}=(D^1, D^2, \cdots)$ is a countable module, there exists in $\{\mathfrak{M}\}$ a sequence of special weight functions $\phi_1, \phi_2, \cdots$ such that*

$$\text{(i)} \qquad \phi_m = \sum_{N=1}^{\infty} s^N \overset{m}{r}_N \left(\sum_{\tau=1}^{s^N} \overset{N}{D}_{\tau\tau} \right);$$

$$\text{(ii)} \qquad 0 \leqq \overset{m}{r}_N \leqq 1, \qquad \lim_{m\to\infty} \overset{m}{r}_N = 1.$$

As in the proof of Theorem 30, we choose the $\epsilon_N \neq 0$ such that the series (27) is uniformly convergent, and construct the special weight functions $\phi_1, \phi_2, \cdots$ mentioned therein.

By Theorem 30 these $\phi_1, \phi_2, \cdots$ have the properties (i), (ii) of our theorem, so we need prove only that they belong to $\{\mathfrak{M}\}$.

As $f(x)$ belongs to $\{\mathfrak{M}\}$, for every finite subset $a_1, \cdots, a_n$ of $\mathfrak{G}$, all $f(xa_1), \cdots, f(xa_n)$ also do, and with them their continuous function

$$\max(|f(xa_1) - f(a_1)|, \cdots, |f(xa_n) - f(a_n)|) = \underset{y=a_1,\cdots,a_n}{\text{l.u.b.}} (|f(xy) - f(y)|).$$

Since $f(z)$ is almost periodic, we can select a sequence of such sets $a_1, \cdots, a_n$, so that this converges uniformly to the translation function

$$t(x) = \underset{y \in \mathfrak{G}}{\text{l.u.b.}} (|f(xy) - f(y)|).$$

Thus $t(x)$, as well as the $\psi_\epsilon(x) = \omega_\epsilon(t(x))$ of Theorem 22, which is a continuous function of $t(x)$, belongs to $\{\mathfrak{M}\}$, and with it

$$\psi(x) = \frac{\psi_\epsilon(x)}{M_x\psi_\epsilon(x)}.$$

The $\chi_1(x)$ of Lemma 2 contains only such representations as occur in the Fourier series of $\psi(x)$; therefore it also belongs to $\{\mathfrak{M}\}$, and with it $\chi(x)$. The same is true of the χ of Theorem 22 and also of the $\chi_1, \chi_2, \cdots$ in the proof of Theorem 30 as well as of the $\psi_1, \psi_2, \cdots$, since these contain only such representations as occur in the Fourier series of the corresponding functions $\chi_1, \chi_2, \cdots$. It follows finally that $\phi_1, \phi_2, \cdots \subset \{\mathfrak{M}\}$, and the proof is complete.

THEOREM 33. (Isolation Theorem.) *Let F be an almost periodic function, with a Fourier expansion*

$$\sum_D \left(s_D \sum_{\rho,\sigma=1}^{s_D} f_{D\rho\sigma} D_{\rho\sigma} \right), \tag{32}$$

and let $\mathfrak{M}$ be any module. There exists an almost periodic function, we shall denote it by $F_{\mathfrak{M}}$, whose Fourier expansion consists of exactly those terms

$$s_D \sum_{\rho,\sigma=1}^{s_D} f_{D\rho\sigma} D_{\rho\sigma} \tag{33}$$

of (32), *for which D is contained in $\mathfrak{M}$. And we have $F_{\mathfrak{M}} \dashv F$.*

The given module $\mathfrak{M}$ need not be enumerable from the outset, but we may

replace it by the enumerable module generated by those representations of $\mathfrak{M}$ which occur in (32). Hence $\mathfrak{M}$ may be assumed to be enumerable. Let it contain the representations $D^1, D^2, \cdots$. With the special weight functions $\phi_1, \phi_2, \cdots$ from Theorem 32 we construct the functions $\phi_1 \times F, \phi_2 \times F, \cdots$. By Theorem 21 the functions $\phi_m \times F$ are all $\dashv\!\mathfrak{z} F$, and by the properties of ϕ_m, the functions $\phi_m \times F$ are formally convergent. By Theorem 28 the sequence $\phi_m \times F$ has a limit function $f_{\mathfrak{M}}$, and by Theorems 24, 4 and 32 the Fourier series of $f_{\mathfrak{M}}$ has the stated form.

Theorem 34. *Let F be an almost periodic function and $\mathfrak{M}_1, \mathfrak{M}_2, \cdots$ a sequence of monotonically increasing modules which in their sum contain all representations occurring in F.*

The sequence of functions

$$F_{\mathfrak{M}_n} \qquad (n = 1, 2, \cdots) \tag{34}$$

converges to F.

The functions (34) converge formally to F, and are all $\dashv\!\mathfrak{z} F$. By Theorem 28, F is the limit of the sequence.

Theorem 35. (Decomposition Theorem.) *Let F be an almost periodic function and let it be possible to divide the representations occurring in F in a sequence of systems $\mathfrak{s}_1, \mathfrak{s}_2, \mathfrak{s}_3, \cdots$, in such a way that for each $k(=1, 2, 3, \cdots)$, the least module containing $\mathfrak{S}(\mathfrak{s}_1, \mathfrak{s}_2, \cdots, \mathfrak{s}_k)$ has no element in common with $\mathfrak{S}(\mathfrak{s}_{k+1}, \mathfrak{s}_{k+2}, \cdots)$. There exists, for each k, an almost periodic function, we shall denote it by $F_{\mathfrak{s}_k}$, whose Fourier expansion consists of exactly those terms* (33) *of* (32), *for which D is contained in $\mathfrak{s}_k$; and the series*

$$F_{\mathfrak{s}_1} + F_{\mathfrak{s}_2} + \cdots$$

converges to F.

If we denote the smallest module containing $\mathfrak{s}_1, \cdots, \mathfrak{s}_k$ by $\mathfrak{M}_k$, the desired function $F_{\mathfrak{s}_k}$ is $F_{\mathfrak{M}_k}$ for $k=1$, and $F_{\mathfrak{M}_k} - F_{\mathfrak{M}_{k-1}}$ for $k \geqq 2$, and our theorem reduces to Theorem 34.

Part IV. Applications to Hilbert space

In this section we shall assume L to be either a Hilbert space $\mathfrak{H}$, or the space $\mathfrak{B}$ of all bounded transformations in $\mathfrak{H}$. We shall consider these spaces with the help of the various topologies which have been discussed in [4], Chapter IV, as well as in the places quoted there. According to whether we consider $\mathfrak{H}$ in the topology based on the neighborhoods $\mathfrak{U}_1$ or $\mathfrak{U}_2$, we shall denote $\mathfrak{H}$ by $\mathfrak{H}_1$ or $\mathfrak{H}_2$ respectively. As for $\mathfrak{B}$, corresponding to the neighborhoods $\mathfrak{U}_3, \mathfrak{U}_4, \mathfrak{U}_5$ we shall denote it by $\mathfrak{B}_3, \mathfrak{B}_4, \mathfrak{B}_5$, respectively.

THEOREM 36. (α) *If $F(x) \in \mathfrak{H}^{\mathfrak{G}}$ is almost periodic in $\mathfrak{H}_1^{\mathfrak{G}}$ it is also almost periodic in $\mathfrak{H}_2^{\mathfrak{G}}$.* ($\beta$) *If $F(x) \in \mathfrak{B}^{\mathfrak{G}}$ is almost periodic in $\mathfrak{B}_3^{\mathfrak{G}}$ it is also almost periodic in $\mathfrak{B}_4^{\mathfrak{G}}$, and* ($\gamma$) *if it is almost periodic in $\mathfrak{B}_4^{\mathfrak{G}}$ it is also almost periodic in $\mathfrak{B}_5^{\mathfrak{G}}$.*

(δ) *$F(x)$ is almost periodic in $\mathfrak{H}_2^{\mathfrak{G}}$ if and only if the numerical function $(F(x), g)$, which is the inner product of $F(x)$ with the constant $g \in \mathfrak{H}$, is almost periodic for every $g \in \mathfrak{H}$.*

(ϵ) *$F(x)$ is almost periodic in $\mathfrak{B}_4^{\mathfrak{G}}$ if and only if $F(x)f$, that is, the value of $F(x)$ operated upon the constant $f \in \mathfrak{H}$, is almost periodic in $\mathfrak{H}_1^{\mathfrak{G}}$ for every $f \in \mathfrak{H}$.*

(ζ) *$F(x)$ is almost periodic in $\mathfrak{B}_5^{\mathfrak{G}}$ if and only if $F(x)f$ is almost periodic in $\mathfrak{H}_2^{\mathfrak{G}}$ for every $f \in \mathfrak{H}$, that is, if and only if the numerical function $(F(x)f, g)$ is almost periodic for every pair $f, g, \in \mathfrak{H}$.*

Ad (α). The topology of $\mathfrak{H}_2$ is weaker than the topology of $\mathfrak{H}_1$. More precisely: to any $U_2 \in \mathfrak{U}_2$ there corresponds a $U_1 \in \mathfrak{U}_1$ such that $U_1 \subset U_2$. Hence, if a set $S \subset \mathfrak{H}$ is totally bounded in $\mathfrak{H}_1$ it is also totally bounded in $\mathfrak{H}_2$. And this proves the proposition, if we remember Definition 1.

Ad (β) and (γ). A similar argument as ad(α).

Ad (δ). Let S be any set of $\mathfrak{H}_b^{\mathfrak{G}}$. If $g \in \mathfrak{H}$ we denote by S_g the set of numerical functions $(F(x), g)$, $F \in S$, and for any finite number of elements $g_1, \cdots, g_n \in \mathfrak{H}$ we denote by $S_{g_1, \ldots, g_n}$ the set of n-dimensional vector functions with components $(F(x), g_1), \cdots, (F(x), g_n)$, $F \in S$. It is easy to see that S is totally bounded in $\mathfrak{H}_2^{\mathfrak{G}}$ if and only if all sets $S_{g_1, \cdots, g_n}$ are totally bounded. On the other hand, we have to prove that S is totally bounded in $\mathfrak{H}_2$ if and only if all S_g are totally bounded. Thus it is sufficient to prove that, for a fixed S, the total boundedness of all sets S_g implies the total boundedness of all sets $S_{g_1, \cdots, g_n}$. This follows from [4], Theorem 9, if we observe that a vector $\{\phi_1, \cdots, \phi_n\}$ may be considered as a continuous function of its components $\phi_1, \cdots, \phi_n$ in the obvious topology of vector spaces.

Ad (ϵ) and (ζ). A similar argument as ad (δ).

THEOREM 37. *Let $F(x)$ be an almost periodic function, its Fourier expansion being (throughout this part, we will write the Fourier coefficients $\alpha_{\rho\sigma}(D)$ without the factors s_D, the degree of D)*

$$F(x) \sim \sum_{D,\rho,\sigma} \alpha_{\rho\sigma}(D) D_{\rho\sigma}(x). \tag{35}$$

If $F(x)$ is in $\mathfrak{H}^{\mathfrak{G}}$, we have $\alpha_{\rho\sigma}(D) \in \mathfrak{H}$, and $(F(x), g)$, for any $g \in \mathfrak{H}$, has the Fourier expansion

(36) $$(F(x), g) \sim \sum_{D,\rho,\sigma} (\alpha_{\rho\sigma}(D), g) D_{\rho\sigma}(x).$$

If $F(x)$ is in $\mathfrak{B}^{\mathfrak{G}}$, we have $\alpha_{\rho\sigma}(D)\epsilon\mathfrak{B}$; for any $f\epsilon\mathfrak{H}$, $F(x)f$ has the Fourier expansion

(37) $$F(x)f \sim \sum_{D,\rho,\sigma} (\alpha_{\rho\sigma}(D)f)\cdot D_{\rho\sigma}(x),$$

and for any $f, g\epsilon\mathfrak{H}$, $(F(x)f, g)$ has the Fourier expansion

(38) $$(F(x)f, g) \sim \sum_{D,\rho,\sigma} (\alpha_{\rho\sigma}(D)f, g)\cdot D_{\rho\sigma}(x).$$

Remembering the definition of $\alpha_{\rho\sigma}(D)$, it is sufficient to prove

(39) $$M_x(F(x), g) = (M_xF(x), g)$$

in the case of (36), and

(40) $$M_x(F(x)f) = (M_xF(x))f$$

in the case (37); (38) follows from their combination.

The mean with respect to the strong topology ($\mathfrak{H}_1$ or $\mathfrak{B}_4$) has, a fortiori, all the properties of the mean in the weak topology ($\mathfrak{H}_2$ or $\mathfrak{B}_5$). Hence, by the uniqueness property of the mean (Theorem 17), it is sufficient to consider the cases of weak topology. In these cases the relations (39), (40) follow by a new application of Theorem 17, since $(M_xF(x), g)$ and $(M_xF(x))f$ have the properties required in this theorem.

THEOREM 38. *If $F(x)$ is an almost periodic function of the class $\mathfrak{B}_\nu^{\mathfrak{G}}$ ($\nu=3$, 4, 5), and $\alpha\epsilon\mathfrak{B}$, then $\alpha F(x)$ and $F(x)\alpha$ [αF and $F\alpha$ are the operational products of F and α] are again almost periodic functions of the same class; and the Fourier expansions of αF and $F\alpha$ may be obtained from the Fourier expansion of $F(x)$ by term-by-term multiplication with α on the right or on the left.*

The first half of the theorem follows from the fact that, for a fixed α, αF and $F\alpha$ are continuous functions of F in each of the topologies $\mathfrak{B}_3$, $\mathfrak{B}_4$, $\mathfrak{B}_5$. The second half (in the case of αF) follows from Theorem 37, Formula (38), if, in the formula of our statement connecting the expansions of αF and F, we apply (38) and then replace F, f, g by $\alpha F, f, g$ in the left-hand member, and by F, f, α^*g (α^* is the adjoint of α) in the right-hand member and compare the results; in the case of $F\alpha$ we replace F, f, g by $F\alpha, f, g$ and by $F, \alpha f, g$ respectively.

THEOREM 39. *Let $F(x)$ be an almost periodic function of* $\mathfrak{B}_5^{\mathfrak{G}}$; *let* (35) *be its Fourier expansion. If $F(x)$ is a representation of the group by means of unitary transformations, that is, if $F(xy)=F(x)F(y)$, $F(1)=1$, $F(x^{-1})=F(x)^*$, then we have the following*:

(i) *$F(x)$ is almost periodic as a function of* $\mathfrak{B}_4^{\mathfrak{G}}$; *moreover, F is the limit, in the $\mathfrak{B}_4$-topology, of the sequence*

$$F_m = \sum_{N=1}^{m}\left(\sum_{\rho,\sigma=1}^{s_{D^N}} \alpha_{\rho\sigma}(D^N) D_{\rho\sigma}^{N}(x)\right). \tag{41}$$

(ii) *The Fourier coefficients have the properties*

$$\alpha_{\rho\sigma}(D)\alpha_{\tau\upsilon}(E) = 0 \qquad \text{if } D \neq E, \tag{42}$$

$$\alpha_{\rho\sigma}(D)\alpha_{\tau\upsilon}(D) = \delta_{\sigma\tau}\alpha_{\rho\upsilon}(D) \qquad [\rho, \sigma, \tau, \upsilon = 1, \cdots, s_D], \tag{43}$$

$$\alpha_{\rho\sigma}(D)^* = \alpha_{\sigma\rho}(D) \qquad [\rho, \sigma = 1, \cdots, s_D]. \tag{44}$$

(iii) *The system of operators $\alpha_{\rho\rho}(D)$ is a resolution of the identity, that is to say, each of them is a projection operator, any two of them are orthogonal, and their sum is the identity.*

Conversely, if a set of elements $\alpha_{\rho\sigma}(D)$ of $\mathfrak{B}$ fulfills the conditions (ii) *and* (iii), *then the series* (35) *is the Fourier expansion of an almost periodic function $F(x)$ with the assigned properties.*

Remark concerning the theorem. Before proving the theorem we want to point out the algebraic aspect of the proposition.

Since $\alpha_{\rho\rho}(D)$ is a projection it corresponds to a closed linear manifold, say $\mathfrak{M}_\rho^D$. Select an orthonormal system determining $\mathfrak{M}_1^D$: $\psi_{11}^D, \psi_{12}^D, \cdots$ (the sequence is empty, finite, or countable). Now, (43) and (44) imply

$$\alpha_{\rho 1}(D)^*\alpha_{\rho 1}(D) = \alpha_{11}(D), \quad \alpha_{\rho 1}(D)\alpha_{\rho 1}(D)^* = \alpha_{\rho\rho}(D),$$

and this means that $\alpha_{\rho 1}(D)$ maps $\mathfrak{M}_1^D$ in a one-to-one and unitary way on $\mathfrak{M}_\rho^D$, while $\alpha_{\rho 1}(D)f\equiv 0$ if f is orthogonal to $\mathfrak{M}_1^D$. Thus, if we define $\psi_{\rho\nu}(D) = \alpha_{\rho 1}(D)\psi_{1\nu}(D)$, the $\psi_{\rho 1}(D), \psi_{\rho 2}(D), \cdots$ determine $\mathfrak{M}_\rho^D$. By (43) we have $\alpha_{\rho\nu}(D)\alpha_{\nu 1}(D) = \alpha_{\rho 1}(D)$, and this implies $\alpha_{\rho\sigma}(D)\psi_{\sigma\nu}(D) = \psi_{\rho\nu}(D)$.

Since, by (iii), $\sum_{D,\rho}\alpha_{\rho\rho}(D) = 1$, the $\mathfrak{M}_\rho^D$ together determine $\mathfrak{H}$, and, since, by (42), (43), (44), $\alpha_{\sigma\rho}(D)^*\alpha_{\tau\upsilon}(E) = 0$ if $D\neq E$ or $\sigma\neq\tau$, the $\mathfrak{M}_\rho^D$ are mutually orthogonal. Thus the $\psi_{\rho\nu}(D)$ form a complete orthonormal system in $\mathfrak{H}$. For this system we found

$$\alpha_{\tau\upsilon}(E)\psi_{\rho\nu}(D) = \begin{cases} 0, \text{ if } D \neq E \text{ or } \upsilon \neq \rho, \\ \psi_{\tau\nu}(D) \text{ if } D = E,\ \upsilon = \rho. \end{cases}$$

Remembering that the $\alpha_{\tau\nu}(D)$ form the Fourier expansion of $F(x)$, we obtain the formulas

$$F(x)\psi_{\rho\nu}(D) = \sum_{\tau=1}^{s_D} D_{\tau\rho}(x)\psi_{\tau\nu}(D).$$

Thus our theorem proves that the representation $F(x)$ may be reduced to split up according to the finite irreducible representations $D(x)$. The subspaces which correspond to this reduction are determined by the

$$\psi_{1\nu}(D),\ \psi_{2\nu}(D),\ \cdots,\ \psi_{s_D\nu}(D) \text{ for all } D,\ \nu.\text{—} \tag{45}$$

We pass on to the proof. Let $F(x)$ have the assumed properties. $F(xy)$ is an almost periodic function in (x, y) and for each y an almost periodic function in x. From the approximating properties of the Fourier expansion it follows that the Fourier expansion of $F_y(x) = F(xy)$ as x-function may be derived from the series (35) in a formal way, namely

$$F_y(x) \sim \sum_{D,\rho,\sigma} \alpha_{\rho\sigma}(D;\ y)D_{\rho\sigma}(x), \tag{46}$$

where

$$\alpha_{\rho\sigma}(D;\ y) = \sum_{\tau=1}^{s_D} \alpha_{\rho\tau}(D)D_{\sigma\tau}(y). \tag{47}$$

On the other hand, $F_y(x) = F(x)F(y)$, and by Theorem 38,

$$F_y(x) \sim \sum \alpha_{\rho\sigma}(D)F(y)D_{\rho\sigma}(x). \tag{48}$$

A comparison of (46) and (48) yields, for any D, ρ, σ,

$$\alpha_{\rho\sigma}(D)F(y) = \sum_{\tau=1}^{s} \alpha_{\rho\tau}(D)D_{\sigma\tau}(y). \tag{49}$$

Now we take y variable, and we again apply Theorem 38. This gives the relations (42) and (43). In order to prove (44) we need only compare the relations

$$F(x^{-1}) \sim \sum \alpha_{\rho\sigma}(D)D_{\rho\sigma}(x^{-1}) = \sum \alpha_{\rho\sigma}(D)\overline{D_{\sigma\rho}(x)}, \tag{50}$$

$$F(x)^* \sim \sum \alpha_{\rho\sigma}(D)^*\overline{D_{\rho\sigma}(x)} \tag{51}$$

(which follow from the approximating properties of the Fourier series), and observe that the representations $\{\overline{D_{\rho\sigma}}\}$ also form a set of irreducible normal representations of $\mathfrak{G}$.

If we apply (44) to $\rho = \sigma$ we find that $\alpha_{\rho\rho}(D)$ is self-adjoint, and (43) gives $(\alpha_{\rho\rho}(D))^2 = \alpha_{\rho\rho}(D)$. Hence, each $\alpha_{\rho\rho}(D)$ is a projection. It follows from (42) and (43) that any two of them are orthogonal. Hence any finite number of

them, and also their (infinite) sum, is a projection again. We still have to prove that $\sum_{D,\rho}\alpha_{\rho\rho}(D)$ is the unity and that (i) holds.

Let f be any element of $\mathfrak{H}$. We want to evaluate the difference

$$F_p(x)f - F_q(x)f, \qquad p > q,$$

$F_p(x)$ being given by (41). Let D be any representation occurring in $F_p(x)f-F_q(x)f$. Writing

$$\beta^D(x) = \sum_{\rho,\sigma=1}^{s_D} \alpha_{\rho\sigma}(D)D_{\rho\sigma}(x),$$

we have

$$\|\beta(x)f\|^2 = (\beta(x)f, \beta(x)f) = \sum_{\rho,\sigma,\tau,\upsilon=1}^{s_D} D_{\rho\sigma}(x)\overline{D_{\tau\upsilon}(x)}(\alpha_{\rho\sigma}(D)f, \alpha_{\tau\upsilon}(D)f);$$

but

$$(\alpha_{\rho\sigma}(D)f, \alpha_{\tau\upsilon}(D)f) = (\alpha_{\tau\upsilon}(D)^*\alpha_{\rho\sigma}(D)f, f) = (\alpha_{\upsilon\tau}(D)\alpha_{\rho\sigma}(D)f, f) = \delta_{\tau\rho}(\alpha_{\upsilon\sigma}(D)f, f);$$

hence

$$\begin{aligned}\|\beta(x)f\|^2 &= \sum_{\upsilon,\sigma=1}^{s_D}\left(\sum_{\tau=1}^{s_D} D_{\tau\sigma}(x)\overline{D_{\tau\upsilon}(x)}\right)(\alpha_{\upsilon\sigma}(D)f, f)\\ &= \sum_{\upsilon,\sigma=1}^{s_D} \delta_{\upsilon\sigma}(\alpha_{\upsilon\sigma}(D)f, f)\\ &= \sum_{\rho=1}^{s_D}(\alpha_{\rho\rho}(D)f, f) = \sum_{\rho=1}^{s_D}\|\alpha_{\rho\rho}(D)f\|^2.\end{aligned}$$

Thus

$$\|F_p(x)f - F_q(x)f\|^2 = \sum_{N=q+1}^{p}\sum_{\rho=1}^{s_{D^N}}\|\alpha_{\rho\rho}(D^N)f\|^2. \tag{52}$$

Since the $\alpha_{\rho\rho}(D)$ are mutually orthogonal projections and the right side of (52) is independent of x, it follows easily that the sequence of functions $F_m(x)$ is uniformly convergent in the topology $\mathfrak{B}_4$. Thus $F(x)$ is the limit of the sequence (41) and is almost periodic also in the class $\mathfrak{B}_4^{\mathfrak{G}}$. Now

$$F(1) = \lim_{m\to\infty} F_m(1) = \lim_{m\to\infty}\sum_{N=1}^{m}\sum_{\rho=1}^{s_{D^N}}\alpha_{\rho\rho}(D^N) = \sum_{D,\rho}\alpha_{\rho\rho}(D).$$

But $F(1)=1$, by assumption, and this proves the last missing part of (iii).

Conversely, let a set of elements $\alpha_{\rho\sigma}(D)$ of $\mathfrak{B}$ satisfy (ii) and (iii). As before, we deduce the relation (52). Hence the sequence $F_m(x)$, defined by

(41), is convergent, in the topology $\mathfrak{B}_4$, to an almost periodic function $F(x)$ whose Fourier expansion is the uniformly convergent series (35). The group properties of $F(x)$ are easily verified from its series (35) on the basis of the known properties (ii), (iii).

THEOREM 40. *Let the group $\mathfrak{G}$ be a metric, locally compact, separable space in which the product xy is a continuous function of (x, y). We consider in $\mathfrak{G}$ a right invariant Haar measure and the Hilbert space $\mathfrak{H}$ consisting of all Lebesgue measurable functions of integrable square in $\mathfrak{G}$, and let $\mathfrak{T}(x)$ denote, for each x, the unitary operator which transforms every element $f(t)\epsilon\mathfrak{H}$ into $f(tx)$.*

In order that $\mathfrak{G}$ be compact it is necessary and sufficient that $\mathfrak{T}(x)$ be an almost periodic function of $\mathfrak{B}_5^{\mathfrak{G}}$.

Obviously $\mathfrak{T}(x)$ has the properties

$$\mathfrak{T}(x)^{-1} = \mathfrak{T}(x)^*, \quad \mathfrak{T}(xy) = \mathfrak{T}(x)\,\mathfrak{T}(y), \quad \mathfrak{T}(1) = 1.$$

It follows easily that $\mathfrak{T}(x)$ is continuous in the topology $\mathfrak{B}_5$.

If $\mathfrak{G}$ is compact, every continuous function is almost periodic and this proves the necessity of our condition. Conversely, let $\mathfrak{T}(x)$ be almost periodic. We consider its Fourier expansion and the complete orthonormal system $\psi_{\rho\nu}(D)$ constructed in the remark to Theorem 39. Consider one of its non-empty subsystems (45) and denote it by $\psi_1, \cdots, \psi_s$. By the last relation of the remark we have

$$F(x)\psi_\rho(t) = \psi_\rho(xt) = \sum_\tau D_{\tau\rho}(x)\psi_\tau(t), \tag{53}$$

except perhaps for a t-set of measure 0, depending on x. By the theorem of Fubini there is a value $t=t_0$ for which (53) holds for all x except a set of measure 0. As $D(x)$ is unitary we obtain

$$\sum_\rho |\psi_\rho(xt_0)|^2 = \sum_\rho |\psi_\rho(t_0)|^2 = C.$$

Upon integrating over $\mathfrak{G}$, the left side gives s_D whereas the right side is C times the total measure of $\mathfrak{G}$. Thus $s_D>0$ implies that $C\neq 0$ and that the total measure of $\mathfrak{G}$ is finite. If the total measure of $\mathfrak{G}$ is finite there cannot exist an $\epsilon_0>0$ and an infinite number of points on $\mathfrak{G}$ any two of which have a distance $>\epsilon_0$. This implies that $\mathfrak{G}$ is totally bounded. Being locally compact, $\mathfrak{G}$ is also complete. Hence the compactness of $\mathfrak{G}$.

BIBLIOGRAPHY

[1] J. von Neumann, *Almost periodic functions in a group*, I, these Transactions, vol. 36 (1934), pp. 445–492.

[2] S. Bochner, *Abstrakte fastperiodische Funktionen*, Acta Mathematica, vol. 61 (1933), pp. 149–184.

[3] S. Bochner, *Fastperiodische Lösungen der Wellengleichung*, Acta Mathematica, vol. 62 (1934), pp. 227–237.

[4] J. von Neumann, *On complete topological spaces*, these Transactions, vol. 37 (1935), pp. 1–20.

[5] S. Bochner, *Beiträge zur Theorie der fastperiodischen Funktionen*, I, Mathematische Annalen, vol. 96 (1926), pp. 119–147.

PRINCETON UNIVERSITY,
PRINCETON, N. J.
INSTITUTE FOR ADVANCED STUDY,
PRINCETON, N. J.

ON REVERSIBLE DYNAMICAL SYSTEMS*

BY
G. BALEY PRICE

Introduction

In recent years great advances have been made in the theory of dynamical systems. Nevertheless comparatively few specific systems have been treated by modern methods in such a way as to set forth the characteristic features of the entire system. Among the specific systems which have been treated, we may mention (1) the restricted problem of three bodies [Birkhoff 5],† which is an irreversible dynamical system and further complicated by the presence of singularities; (2) the determination of the geodesics on a surface of negative curvature [Morse 2, 3; Birkhoff 4, pp. 238–248]; (3) a simple type of reversible dynamical systems on surfaces of revolution [Price 1, 2]; (4) the billiard ball problem [Birkhoff 6; 4, pp. 169–179]. The purpose of the present paper is to study reversible dynamical systems, with the emphasis on those which have an oval of zero velocity.

It has been known for a long time that among the small oscillations [Whittaker 1, chap. VII] of a reversible dynamical system at a position of stable equilibrium there are two fundamental periodic orbits which join two points of the oval of zero velocity and are traced with a backward and forward motion. Part I of the present paper is devoted to proving that similar orbits exist in the actual system at least for sufficiently restricted values of the energy constant. These orbits appear to be one of the characteristic features of a reversible dynamical system with an oval of zero velocity. The property of reversible systems which distinguishes them from irreversible systems is that their trajectories in the manifold of states of motion are grouped in symmetric pairs [see §2].

Using methods developed by Poincaré [1, vol. I, chap. III] and elaborated by Painlevé [1], Horn [1] has shown that there exist certain periodic orbits in the actual system which reduce to the fundamental periodic orbits of the small oscillations system, but their exact nature was not determined. In the present treatment a parameter μ is introduced in such a way that $\mu=0$

---

* Part of a paper presented to the Society, February 25, 1933, under the title *A study of certain dynamical systems with applications to the generalized double pendulum*; received by the editors February 9, 1934.

† We shall refer in this manner to the bibliography at the end of this paper.

gives the limiting case of small oscillations. Then the equations of motion are integrated in terms of series in parameters by Poincaré's method, and two theorems on the analytic continuation of the fundamental periodic orbits of the limiting integrable system are proved. The second theorem applies only when the system is symmetric in the position of equilibrium. These theorems show not only that analytic continuation is possible in certain cases in which Horn's method fails, but also that the continued orbits always touch the oval of zero velocity.

Part II, using the results of Part I, is devoted to a detailed study of motion in the neighborhood of a position of stable equilibrium in the case of two degrees of freedom. The methods are those of analytic continuation [Birkhoff 4, pp. 139–143; Poincaré 1, vol. I, chap. III], surfaces of section, and surface transformations [Birkhoff 1, 3]. First, the manifold of states of motion is studied, and a convenient representation of it is given in 3-space. Then the limiting integrable system is treated in detail. Next some general theorems on a common type of surface of section are given. At this point the presence of the oval of zero velocity introduces essential difficulties, since the surface of section is formed from a periodic orbit which joins two points of it. The existence of periodic orbits is established by applying the general theory of surface transformations. Poincaré's Last Geometric Theorem is useful here. When the system is symmetric in the position of equilibrium, the transformation on the surface of section can be factored in certain ways, and more specific results concerning periodic orbits are established.

The author wishes to acknowledge his indebtedness and express his thanks to Professor Birkhoff, who proposed a problem which led to this paper and made valuable suggestions concerning the method of treatment.

Part I. Analytic continuation of periodic orbits

1. Definitions and assumptions. Let (y_i) and (y_i'), $i=1, \cdots, n$, be the position and velocity coordinates respectively of a dynamical system with kinetic energy T and force function U. Throughout the paper a prime will denote a derivative with respect to the time. Let the (y_i) be principal coordinates [Whittaker 1, chap. VII], and assume that T and U have the following specific forms:

$$(1)\qquad T = \frac{1}{2}\sum_{i,j=1}^{n}[\delta_{ij} + T_{ij}(y_1, \cdots, y_n)]y_i' y_j' \qquad (\delta_{ii} = 1;\ \delta_{ij} = 0,\ i \neq j),$$

$$(2)\qquad U = \frac{1}{2}\sum_{i=1}^{n}\lambda_i^2 y_i^2 + u(y_1, \cdots, y_n),$$

where

$$(3) \qquad | \delta_{ij} + T_{ij}(y_1, \cdots, y_n) | \neq 0,$$

and the λ_i are either real or pure imaginary with

$$(4) \qquad \lambda_i \neq 0 \qquad (i = 1, \cdots, n).$$

Here the $T_{ij}(y_i)$ are functions which vanish at (0), and $u(y_i)$ has no terms of degree lower than the third; we assume that these are entire functions of the indicated arguments.

The equations of motion in the Lagrangian form can be written down at once; because the determinant in (3) is not zero, these equations can be put in the form

$$(5) \qquad y_i'' = \lambda_i^2 y_i + \sum_{k,j=1}^{n} F_{kj}^{(i)}(y_1, \cdots, y_n) y_k' y_j' + G_i(y_1, \cdots, y_n),$$

where $i = 1, \cdots, n$. The integral of energy is

$$(6) \qquad T = U + h/2.$$

Since we assume T to be a positive definite quadratic form in the velocities, it follows from (6) that the motion takes place in the regions $U + h/2 \geqq 0$, bounded by the *oval of zero velocity* Z with the equation $U + h/2 = 0$. This system may be interpreted as a particle of unit mass moving on a surface with

$$(7) \qquad ds^2 = 2T(dt)^2$$

and (y_i) as the coordinates, and acted on by forces derived from U [Birkhoff 1, pp. 202 and 212–213; 4, pp. 23–25]. A curve on the *characteristic surface* (7) defined by a solution $[y_i(t)]$ of (5) will be termed an *orbit* of the particle, and the curve $[y_i(t); y_i'(t)]$ in the *manifold of states of motion M* with the equation (6) will be called a *trajectory* or *stream line*.

2. **Reversible dynamical systems.** Now M is a $(2n-1)$-manifold in the $2n$-space with coordinates $(y_i; y_i')$. We agree once and for all to exclude those values of h which lead to a double point on Z, i.e., we assume the equations

$$(8) \qquad U + h/2 = 0, \qquad \partial U/\partial y_i = 0 \qquad (i = 1, \cdots, n)$$

have no simultaneous solution. With this restriction, M is an analytic manifold, and there is no equilibrium solution $y_i = \text{constant}$.

The usual existence theorems of differential equations [Birkhoff 4, pp. 1–14] can be applied to the system (5), (6). The orbits are regular curves except at those points at which they touch Z, for only at these points can (y_i') vanish. If an orbit touches Z, the particle approaches and recedes from

it along one and the same curve, and spends only a finite length of time in the neighborhood of the point. On the other hand, the trajectories are the stream lines of a steady fluid motion in M. Since there are no double points on Z, $(y_i';y_i'')$ cannot vanish; hence the trajectories are regular curves without exception.

Dynamical systems of the type introduced in §1 are known as reversible [Birkhoff 1, p. 205; 4, pp. 27–29]; their fundamental property is stated in the following theorem.

THEOREM 1. *An orbit of a reversible dynamical system may be traced in either direction; or again, the stream lines are paired, each stream line of a pair being the symmetric image of the other in the n-plane* $y_i'=0$, $i=1,\cdots,n$.

The proof follows from the fact that if $[y_i(t);\,y_i'(t)]$ is a solution of (5), then $[y_i(-t);\,-y_i'(-t)]$ is also a solution. The coordinates of symmetric points on the two trajectories are obtained by combining the coordinates of a point on the orbit with the two possible directions of the velocity vector. Each trajectory of a pair will be called the *symmetric trajectory of the first kind* of the other. If T^* is a trajectory, its symmetric image of the first kind may be represented by V_1T^*, where V_1 may be thought of as a transformation.

THEOREM 2. *A necessary and sufficient condition that* T^* *and* V_1T^* *be identical is that the orbit which corresponds to* T^* *pass through a point of* Z.

The proof of this theorem and the following one are left to the reader.

THEOREM 3. *A necessary and sufficient condition that an orbit which passes through a point of* Z *be periodic is that it pass through a second point of* Z, *distinct from the first.*

A periodic orbit which passes through a point of Z is thus a curve joining two points of Z, and the particle traces this curve with a backward and forward motion. The length of time required for the particle to pass from one of the points of Z to the other is the same in either direction.

In certain cases, T and U are symmetric in the origin of coordinates on the characteristic surface, i.e.,

$$T(-y_i;\,y_i') = T(y_i;\,y_i'),\qquad U(-y_i) = U(y_i). \tag{9}$$

THEOREM 4. *If* T *and* U *satisfy* (9), *the orbits are paired, each orbit of a pair being the symmetric image of the other in the origin; or again, the trajectories are paired, each trajectory of a pair being the symmetric image of the other in the origin of coordinates in* M.

The proof follows from the fact that if $[y_i(t);\,y_i'(t)]$ is a trajectory, then

$[-y_i(t);\ -y_i'(t)]$ is also a trajectory when (9) holds. Each trajectory of a pair will be called the *symmetric trajectory of the second kind* of the other. If T^* is a trajectory, then V_2T^* will denote the symmetric trajectory of the second kind. Thus when (9) holds, the trajectories are related by fours. The group associated with T^* is T^*, V_2T^*, V_1T^*, and $V_1V_2T^*$.

THEOREM 5. *If T and U satisfy* (9), *a necessary and sufficient condition that an orbit which passes through the origin be periodic is that it pass through the origin a second time.*

The proof follows from the fact that if an orbit passes through the origin, it is its own symmetric image in the origin. Hence, the complete orbit can be obtained by reflecting in the origin the part between any two successive passages through the origin. If an orbit passes through the origin at time $t=t_0$ and closes at time $t=t_1$, then it passes through the origin at time $t=(t_0+t_1)/2$ also.

As a consequence of Theorems 1 and 5, we have the following theorem, of special importance later.

THEOREM 6. *If an orbit passes through the origin at time $t=0$ and touches Z at time $t=t^*$, it is a periodic orbit with period $4t^*$ and joins two points of Z which are symmetric in the origin.*

3. **The limiting integrable system.** The equations of motion and the integral of energy are given by (5) and (6). We now replace the energy constant h by μ^2 and make the change of variables

$$y_i = \mu x_i \qquad (i = 1, \cdots, n). \tag{10}$$

The equations of motion and the integral of energy in the new variables are

$$\begin{gathered} x_i'' = \lambda_i^2 x_i + \mu \sum_{k,j=1}^{n} F_{kj}^{(i)}(\mu x_1, \cdots, \mu x_n) x_k' x_j' + \frac{1}{\mu} G_i(\mu x_1, \cdots, \mu x_n), \\ \sum_{i=1}^{n} x_i'^2 + \sum_{k,j=1}^{n} T_{kj}(\mu x_1, \cdots, \mu x_n) x_k' x_j' \\ = \sum_{i=1}^{n} \lambda_i^2 x_i^2 + \frac{1}{\mu^2} u(\mu x_1, \cdots, \mu x_n) + 1, \end{gathered} \tag{11}$$

where $i=1, \cdots, n$, and T_{kj}, G_i, and u are entire functions whose series expansions have no terms of degree lower than the first, second, and third respectively. Now for every value of $\mu \neq 0$ these equations represent the actual system (5) and (6) for the value $h=\mu^2$ of the energy constant. On the other hand, μ may be considered as a parameter in the equations of the

system. The system is analytic in the parameter for all values of the parameter, including $\mu=0$.

For $\mu=0$ the system (11) is integrable, being

$$x_i'' = \lambda_i^2 x_i \qquad (i = 1, \cdots, n),$$
$$\sum_{i=1}^{n} x_i'^2 = \sum_{i=1}^{n} \lambda_i^2 x_i^2 + 1. \tag{12}$$

This system is a limiting case of the actual system (5) and (6) obtained by reducing the energy to zero and at the same time altering the units of length according to (10). The solution of (12) for which $(x_i; x_i')$ reduces at time $t=0$ to $(\alpha_i; \beta_i)$ is

$$\begin{aligned} x_i &= (\alpha_i/2)[\exp(\lambda_i t) + \exp(-\lambda_i t)] + (\beta_i/(2\lambda_i))[\exp(\lambda_i t) - \exp(-\lambda_i t)], \\ x_i' &= (\alpha_i\lambda_i/2)[\exp(\lambda_i t) - \exp(-\lambda_i t)] + (\beta_i/2)[\exp(\lambda_i t) + \exp(-\lambda_i t)], \end{aligned} \tag{13}$$

where $i=1, \cdots, n$, and

$$\sum_{i=1}^{n} \beta_i^2 = \sum_{i=1}^{n} \lambda_i^2 \alpha_i^2 + 1. \tag{14}$$

We now suppose that k of the λ_i, $0<k\leqq n$, are pure imaginary. We can suppose the notation is so chosen that they are $\lambda_1, \cdots, \lambda_k$. Then among the solutions (13) there are k of special importance.

THEOREM 7. *The limiting integrable system* (12) *has the k fundamental periodic trajectories*

$$\begin{aligned} x_i &= 0, & \\ x_i' &= 0 & (i\neq j;\ i = 1, \cdots, n), \\ x_j &= (1/|\lambda_j|)\sin(|\lambda_j| t + \theta_j), & \\ x_j' &= \qquad\quad \cos(|\lambda_j| t + \theta_j) & (j = 1, \cdots, k). \end{aligned} \tag{15}$$

4. **Solutions of the equations of motion in terms of series in parameters.** Our ultimate aim is to show that analytic continuation of the fundamental periodic trajectories of the limiting integrable system is possible. With this end in view, we shall obtain the solution of (11) in terms of series in certain parameters, following a method developed by Poincaré [1, vol. I, pp. 58–63; Moulton 1, chap. III].

Set

$$x_i = x_i, \qquad x_i' = y_i \qquad (i = 1, \cdots, n). \tag{16}$$

The system (11) thus takes the form

$$
(17)\qquad
\begin{aligned}
&\frac{dx_i}{dt} = y_i,\\
&\frac{dy_i}{dt} = \lambda_i^2 x_i + \mu Y_i(x_1, \cdots, x_n; y_1, \cdots, y_n; \mu),\\
&\sum_{i=1}^{n} y_i^2 + \sum_{i,j=1}^{n} T_{ij}(\mu x_1, \cdots, \mu x_n) y_i y_j\\
&\qquad = \sum_{i=1}^{n} \lambda_i^2 x_i^2 + \frac{1}{\mu^2} u(\mu x_1, \cdots, \mu x_n) + 1,
\end{aligned}
$$

where $i=1, \cdots, n$, and the Y_i are entire functions of the indicated arguments. Now set

$$
(18)\qquad t = t^*(t_0 + \tau)/t_0.
$$

Here t^* is the new independent variable, τ is a parameter, and t_0 is a constant whose value will be specified later. Now transform from the variables $(x_i; y_i)$ to new variables $(p_i; q_i)$ by means of

$$
(19)\qquad
\begin{aligned}
&x_i = p_i + \gamma_i,\\
&y_i = q_i + \delta_i \qquad\qquad (i \neq j,\ i = 1, \cdots, n),\\
&x_j = p_j + \gamma_j + \frac{1}{|\lambda_j|} \sin\left(\frac{t_0+\tau}{t_0} |\lambda_j| t^* + \theta_j\right),\\
&y_j = q_j + \delta_j + \qquad \cos\left(\frac{t_0+\tau}{t_0} |\lambda_j| t^* + \theta_j\right).
\end{aligned}
$$

Here $(\gamma_i; \delta_i)$ are to be considered as parameters in the transformation. The equations of motion in the new variables are

$$
(20)\qquad
\begin{aligned}
&\frac{dp_i}{dt^*} = \frac{t_0+\tau}{t_0}(q_i + \delta_i),\\
&\frac{dq_i}{dt^*} = \frac{t_0+\tau}{t_0}\left[\lambda_i^2(p_i + \gamma_i) + \mu Q_i(p_i; q_i; \gamma_i; \delta_i; \mu; \tau; t^*)\right],
\end{aligned}
$$

where $i=1, \cdots, n$. The right hand members of these equations can be expanded as power series in the $2n$ variables $(p_i; q_i)$ and the $(2n+2)$ parameters $(\gamma_i; \delta_i; \mu; \tau)$ with coefficients which are analytic functions of t^*. On carrying through the details of Poincaré's method and transforming back to the original variables $(x_i; y_i)$, we find the following solution of (17):

$$
(21)\quad
\begin{aligned}
x_i &= (1/2)[\exp(\lambda_i t^*) + \exp(-\lambda_i t^*)]\gamma_i \\
&\quad + (1/(2\lambda_i))[\exp(\lambda_i t^*) - \exp(-\lambda_i t^*)]\delta_i + \cdots, \\
y_i &= (\lambda_i/2)[\exp(\lambda_i t^*) - \exp(-\lambda_i t^*)]\gamma_i \\
&\quad + (1/2)[\exp(\lambda_i t^*) + \exp(-\lambda_i t^*)]\delta_i + \cdots, \\
x_j &= \frac{1}{|\lambda_j|}\sin\left[\frac{t_0+\tau}{t_0}|\lambda_j|\,t^* + \theta_j\right] + (1/2)[\exp(\lambda_j t^*) + \exp(-\lambda_j t^*)]\gamma_j \\
&\quad + (1/(2\lambda_j))[\exp(\lambda_j t^*) - \exp(-\lambda_j t^*)]\delta_j + \cdots, \\
y_j &= \cos\left[\frac{t_0+\tau}{t_0}|\lambda_j|\,t^* + \theta_j\right] + (\lambda_j/2)[\exp(\lambda_j t^*) - \exp(-\lambda_j t^*)]\gamma_j \\
&\quad + (1/2)[\exp(\lambda_j t^*) + \exp(-\lambda_j t^*)]\delta_j + \cdots,
\end{aligned}
$$

where $i \neq j$, $i=1, \cdots, n$. This solution has the following properties:

(I) The series in (21) are series in the $(2n+2)$ parameters $(\gamma_i;\ \delta_i;\ \mu;\ \tau)$. Except for the terms in μ, the series are written out completely up to terms of the second degree. The coefficients are real analytic functions of the real variable t^*. From (18) we see that the values of $(x_i;\ y_i)$ at time $t=t_0+\tau$ are obtained from (21) by setting $t^*=t_0$.

(II) If T^* be chosen arbitrarily, it is possible to find an ϵ such that the series converge absolutely and uniformly for $0 \leqq t^* \leqq T^*$, $|\gamma_i| \leqq \epsilon$, $|\delta_i| \leqq \epsilon$, $|\mu| \leqq \epsilon$, $|\tau| \leqq \epsilon$.

(III) The coefficients of all terms not explicitly written out in (21) vanish for $t^*=0$. The solution (21) thus satisfies the initial conditions

$$
(22)\quad
\begin{aligned}
\alpha_i &= \gamma_i, \\
\beta_i &= \delta_i \qquad\qquad (i \neq j,\ i = 1, \cdots, n), \\
\alpha_j &= (1/|\lambda_j|)\sin\theta_j + \gamma_j, \\
\beta_j &= \qquad\quad \cos\theta_j + \delta_j.
\end{aligned}
$$

(IV) For $\gamma_i=\delta_i=\mu=\tau=0$, the trajectory (21) reduces to the fundamental periodic trajectory (15) of the limiting integrable system.

5. **Analytic continuation of the fundamental periodic trajectories of the limiting integrable system.** We shall now show that under certain conditions each of the k periodic trajectories (15) can be continued analytically for $\mu > 0$. We shall give the proof in the case $j=1$; the proof in the other cases is similar.

From the solutions (21) we select for special consideration those for which $\theta_j=\pi/2$, $\delta_i=0$, $i=1, \cdots, n$. They are

$$
(23)\quad
\begin{aligned}
x_1 &= (1/|\lambda_1|)\cos[(t_0+\tau)|\lambda_1|t^*/t_0] + (1/2)[\exp(\lambda_1 t^*) + \exp(-\lambda_1 t^*)]\gamma_1 \\
&\qquad + \cdots, \\
y_1 &= \qquad -\sin[(t_0+\tau)|\lambda_1|t^*/t_0] + (\lambda_1/2)[\exp(\lambda_1 t^*) - \exp(-\lambda_1 t^*)]\gamma_1 \\
&\qquad + \cdots, \\
x_i &= \qquad (1/2)[\exp(\lambda_i t^*) + \exp(-\lambda_i t^*)]\gamma_i \\
&\qquad + \cdots, \\
y_i &= \qquad (\lambda_i/2)[\exp(\lambda_i t^*) - \exp(-\lambda_i t^*)]\gamma_i \\
&\qquad + \cdots,
\end{aligned}
$$

where $i=2, \cdots, n$. The orbits corresponding to (23) are characterized by the fact that at time $t^*=0$ they pass through a point of Z. Now when $\mu=0$, it is possible to determine the parameters γ_i and τ so that the orbit (23) passes through a second point of Z at time $t^*=t_0=\pi/|\lambda_1|$. We hereby define t_0 in (18). For $\mu=0$, we have only to take $\gamma_i=0$, $\tau=0$.

We now seek to determine $(\gamma_i;\tau)$ as functions of μ so that for every value of μ the orbit (23) passes through a second point of Z at time $t^*=t_0=\pi/|\lambda_1|$. A necessary and sufficient condition that an orbit pass through a point of Z is that all the velocities vanish simultaneously. We therefore have the following equations for determining $(\gamma_i;\tau)$ as functions of μ:

$$
(24)\quad
\begin{aligned}
-\sin(|\lambda_1|\tau+\pi) - |\lambda_1|(\sin\pi)\gamma_1 + \cdots &= 0, \\
-|\lambda_i|\left(\sin\left|\frac{\lambda_i}{\lambda_1}\right|\pi\right)\gamma_i + \cdots &= 0, \\
\frac{\lambda_j}{2}\left[\exp\left(\frac{\lambda_j\pi}{|\lambda_1|}\right) - \exp\left(-\frac{\lambda_j\pi}{|\lambda_1|}\right)\right]\gamma_j + \cdots &= 0, \\
\lambda_1^2\left(\frac{1}{|\lambda_1|}+\gamma_1\right)^2 + \sum_{i=2}^{n}\lambda_i^2\gamma_i^2 + \frac{1}{\mu^2}u(\mu\gamma_1, \cdots, \mu\gamma_n) + 1 &= 0.
\end{aligned}
$$

Here $i=2, \cdots, k$ and $j=k+1, \cdots, n$. The last equation states that the integral of energy is satisfied at time $t^*=0$. The equations (24) are power series in $(\gamma_i;\tau;\mu)$ with constant coefficients. For $\mu=0$ they have the solution $\gamma_i=\tau=0$. If the Jacobian with respect to $(\gamma_i;\tau)$ is not zero for $\gamma_i=\tau=0$, it is possible to solve (24) and obtain $(\gamma_i;\tau)$ as analytic functions of μ at least for μ small. Direct computation shows that this Jacobian does not vanish unless $|\lambda_i|/|\lambda_1|$, $i=2, \cdots, k$, is an integer. Since an orbit which passes through two distinct points of Z is periodic by Theorem 3, we have proved the following theorem [compare Birkhoff 4, pp. 139–143].

THEOREM 8. *The jth fundamental periodic orbit of the limiting integrable system can be continued analytically for $\mu>0$, at least for μ small, if λ_i/λ_j, $i\neq j$, $i=1, \cdots, k$, is not an integer; each orbit of the continuation joins two points of Z and is periodic with period $2(\pi/|\lambda_j|+\tau)$.*

If the system has the symmetry specified by (9), another procedure is possible, which gives additional information and in certain cases additional results. Consider the trajectories (21) for which

$$\theta_i = 0, \qquad \gamma_i = 0 \qquad (i = 1, \cdots, n). \tag{25}$$

The characteristic property of the corresponding orbits is that they pass through the origin when $t^*=0$. By setting $\delta_i=\tau=0$ when $\mu=0$, we obtain an orbit which passes through a point of Z when $t^*=t_0=\pi/(2|\lambda_1|)$. We hereby define t_0 anew in (18). We propose to show that under certain conditions it is possible to determine $(\delta_i; \tau)$ as functions of μ so that the orbit determined by (21) and (25) has this property for all values of μ sufficiently small. The following equations determine these functions:

$$\begin{aligned} \cos(\pi/2 + |\lambda_1|\tau) + (\cos\pi/2)\delta_1 + \cdots &= 0, \\ \left(\cos\left|\frac{\lambda_i}{\lambda_1}\right|\frac{\pi}{2}\right)\delta_i + \cdots &= 0, \\ (1/2)\left[\exp\left(\frac{\pi\lambda_j}{2|\lambda_1|}\right) + \exp\left(-\frac{\pi\lambda_j}{2|\lambda_1|}\right)\right]\delta_j + \cdots &= 0, \\ (1+\delta_1)^2 + \delta_2^2 + \cdots + \delta_n^2 - 1 &= 0. \end{aligned} \tag{26}$$

Here $i=2, \cdots, k$, and $j=k+1, \cdots, n$. The first n equations express the condition that the velocities vanish for $t^*=t_0=\pi/(2|\lambda_1|)$, and the last equation states that the initial conditions satisfy the integral of energy. The equations (26) determine the analytic continuation of the first fundamental periodic trajectory (15); the equations for the others are similar.

Equations (26) have the solution $\delta_i=\tau=0$ when $\mu=0$. We can solve and get $(\delta_i; \tau)$ as analytic functions of μ if the Jacobian does not vanish when $\delta_i=\tau=0$. A direct computation shows that this Jacobian vanishes if and only if λ_i/λ_1, $i=2, \cdots, k$, is an *odd* integer. The orbits (21) for which (9), (25), and (26) hold pass through the origin for $t^*=0$ and touch Z when $t^*=\pi/(2|\lambda_1|)$; hence, by Theorem 6 they are periodic. We have thus proved the following theorem.

THEOREM 9. *If (9) holds, and if λ_i/λ_j, $i\neq j$, $i=1, \cdots, k$, is not an odd integer, the jth fundamental periodic orbit of the limiting integrable system can be continued analytically for $\mu>0$; each orbit of the continuation is a curve which*

passes through the origin, is symmetric in the origin, joins two points of Z, and is periodic with period $4(\pi/(2|\lambda_j|)+\tau)$.

If the system is symmetric in the origin, the periodic orbits whose existence is established by Theorem 8 are identical with those established by Theorem 9, because the continuation in each case is unique. However, since Theorem 9 fails only when λ_i/λ_j is an odd integer, we see that it proves that analytic continuation is possible in some cases when the first theorem fails.

COROLLARY 1. *At a maximum of the force function U in the case of two degrees of freedom at least one of the fundamental periodic orbits of the limiting integrable system can be continued analytically for* $\mu>0$ *unless* $\lambda_1=\lambda_2$.

This corollary follows from Theorem 8 and the fact that λ_1/λ_2 and λ_2/λ_1 are not both integers unless $\lambda_1=\lambda_2$. Important use will be made of this corollary in later work.

PART II. MOTION IN THE NEIGHBORHOOD OF A POSITION OF STABLE EQUILIBRIUM IN THE CASE OF TWO DEGREES OF FREEDOM

6. **The manifold of states of motion.** We continue the study of the dynamical systems of Part I, but we now restrict attention to motion in the neighborhood of a position of stable equilibrium in the case of two degrees of freedom. First we shall investigate the manifold of states of motion M.

The equation of M is $T=U+h/2$. We restrict h henceforth to values for which the region of motion R about the origin on the characteristic surface is homeomorphic to a circular disc. The oval of zero velocity is a simple closed curve Z. There may be other regions of motion for the given value of h, but attention will be confined to the one R about the origin.

Suppose first that h is so restricted that U has only a single critical point in R, a maximum at the origin [see (2) and (4)]. Then the contour curves $U+h/2=c$ are simple closed curves surrounding the critical point. Then a homeomorphism between the points of R and the unit circle C: $u^2+v^2\leqq 1$ can be established as follows. Let the points on the curves $U+h/2=c$ correspond in a one-to-one and continuous manner with the points on the circle $u^2+v^2=(1-2c/h)$, each point (x_1^0, x_2^0) of R† corresponding to a point (u_0, v_0) of C. Then as c varies from $h/2$ to 0, the contour curve expands from the origin and sweeps through R; the corresponding circle expands from the origin and sweeps through C. Now a point of M is obtained by combining the coordinates (x_1^0, x_2^0) of a point of R with the coordinates of a point (y_1, y_2) on the ellipse $T(x_1^0, x_2^0; y_1, y_2)=U(x_1^0, x_2^0)+h/2$ [see (16) for the notation]. This ellipse is real and non-degenerate if (x_1^0, x_2^0) is an interior point of R; it is the

† For convenience, the variables y of §§1 and 2 have been replaced by x.

point ellipse $y_1=y_2=0$ when (x_1^0, x_2^0) is on Z. Let the points on this ellipse correspond to the points on the circle $\xi^2+\eta^2=1-(u_0^2+v_0^2)$, the points on corresponding rays through the origins corresponding. This circle degenerates to a point when and only when the ellipse degenerates to a point. We thus establish a one-to-one and continuous correspondence between the points of M and the unit 3-sphere S_3: $u^2+v^2+\xi^2+\eta^2=1$ in 4-space.

It is possible to give a representation of M in 3-space. Put R into correspondence with C in the way explained above. Then put the points of the ellipse $T(x_1^0, x_2^0; y_1, y_2)=U(x_1^0, x_2^0)+h/2$ into one-to-one and continuous correspondence with the points of the line segment $-[1-(u_0^2+v_0^2)]^{1/2} \leqq w \leqq [1-(u_0^2+v_0^2)]^{1/2}$, the two end points being considered identical, by means of

$$(27)\qquad (\text{arc tan } y_2/y_1)[1-(u_0^2+v_0^2)]^{1/2}=\pi w, \qquad -\pi \leqq \text{arc tan } y_2/y_1 \leqq \pi.$$

By definition, w shall be zero when y_1 and y_2 vanish simultaneously. We have thus put M into one-to-one and continuous correspondence with the points of the unit sphere S_2: $u^2+v^2+w^2\leqq 1$, the points (u, v, w) and $(u, v, -w)$ of the bounding sphere being considered identical.

Consider the general case now. Assume that R is homeomorphic to C with no restriction on the number of critical points of U. Then by the method just explained, we can put M into correspondence with S_2 with the stated convention about points of the bounding sphere. But S_2 can be put into one-to-one and continuous correspondence with S_3, the unit 3-sphere in 4-space. We have proved this, because in the first case we put both into correspondence with M. We have thus proved the following theorem.

THEOREM 10. *If R is homeomorphic to a circular disc, then M is homeomorphic to S_3, and also to S_2 with the points (u, v, w) and $(u, v, -w)$ of the bounding sphere considered identical.*

It is obvious how these results are to be extended to dynamical systems with n degrees of freedom.

We have also the following important theorem concerning the steady fluid motion in M.

THEOREM 11. *The steady flow in M possesses an invariant volume integral.*

This result may be established most easily by transforming to Hamiltonian coordinates [Birkhoff 4, p. 212]. The result is well known, and the details are omitted [see also Birkhoff 1, pp. 211–212; Poincaré 1, vol. III, chaps. XXII–XXIII].

7. **The limiting integrable system.** A detailed study of the limiting integrable system will be made now. By setting $n=k=2$ in §3, we find that the equations of motion are

$$\frac{dx_i}{dt} = y_i, \qquad \frac{dy_i}{dt} = \lambda_i^2 x_i \qquad (i = 1, 2) \tag{28}$$

and that the integral of energy is

$$\sum_{i=1}^{2} (y_i^2 + |\lambda_i|^2 x_i^2) = 1. \tag{29}$$

The general solution of (28) and (29) is

$$\begin{aligned} x_i &= \frac{a_i}{|\lambda_i|} \sin (|\lambda_i| t + \theta_i), \qquad a_1^2 + a_2^2 = 1, \\ y_i &= a_i \cos (|\lambda_i| t + \theta_i) \qquad (i = 1, 2). \end{aligned} \tag{30}$$

The region of motion R^0 is the interior of the ellipse $|\lambda_1|^2 x_1^2 + |\lambda_2|^2 x_2^2 = 1$, whose boundary is the oval of zero velocity Z^0. The axes of this ellipse lie along the x_1- and x_2-axes and are respectively the first and second fundamental periodic orbits O_1^0 and O_2^0. The manifold of states of motion M^0 is the ellipsoid (29).

Consider also the representation of M^0 in S_2. Now R^0 is mapped on C by

$$u = |\lambda_1| x_1, \qquad v = |\lambda_2| x_2. \tag{31}$$

The representation in S_2 can be completed as explained in §6. A trajectory corresponds to a curve in S_2 which may be called a stream line or line of flow. Consider in particular the lines of flow which represent O_1^0 and O_2^0. From (27) and (31) we see that O_2^0 is represented by the ellipse

$$v^2 + 4w^2 = 1, \ u = 0. \tag{32}$$

The direction of flow is the same as that of the rotation which carries the positive w-axis into the positive v-axis. Similarly, O_1^0 is represented by a curve in $v=0$. It is composed of the diameter of S_2 which lies along the u-axis and the semi-circle $w=(1-u^2)^{1/2}$, $v=0$ [or $w=-(1-u^2)^{1/2}$, $v=0$]. The flow is such that its direction is positive along the u-axis.

THEOREM 12. *The surface SS^0: $x_1=0$, $y_1 \geqq 0$ is a surface of section for the limiting integrable system.*

From (29) we see that SS^0 is the semi-ellipsoid

$$y_1^2 + y_2^2 + |\lambda_2|^2 x_2^2 = 1, \ y_1 \geqq 0, \tag{33}$$

in the plane $x_1=0$. The boundary is given by $x_1=0$, $y_1=0$; it is the ellipse which bounds the semi-ellipsoid. The boundary of SS^0 is therefore the closed stream line which corresponds to O_2^0. Equations (30) show that any stream line crosses SS^0 when t has a value which satisfies $|\lambda_1|t+\theta_1=2m\pi$, m any integer; hence, every stream line crosses SS^0 an infinite number of times, and the interval of time between any two successive crossings is $2\pi/|\lambda_1|$.

Next we must show that the angle at which a trajectory crosses SS^0 is of the first order in the distance to the boundary. The direction components of the stream line are given by (28); the surface of section is defined as the intersection of the two 3-spaces $\sum_{i=1}^{2}(y_i^2+|\lambda_i|^2x_i^2)=1$, $x_1=0$ with $y_1\geqq 0$. By a straightforward calculation, using the formula developed in the next section for the angle of intersection of a curve and a 2-surface in 4-space, we find that if ψ is the angle at which a stream line crosses SS^0, then

$$\sin\psi = \frac{y_1}{(\lambda_2^4x_2^2+y_1^2+y_2^2)^{1/2}}. \tag{34}$$

Since SS^0 is the semi-ellipsoid (33), it is clear that y_1 may be taken as a measure of the distance of a point on it to the boundary. From (34) it then follows that ψ is of the first order in the distance to the boundary. The fact that every crossing of SS^0 by a stream line is in the same sense follows from $\sin\psi\geqq 0$ in (34), but it will be geometrically obvious when we consider the representation of SS^0 in S_2. Thus SS^0 satisfies all the requirements of Birkhoff's definition of a surface of section [Birkhoff 1, p. 268], and the proof is complete.

Now consider the representation of SS^0 in S_2. Since SS^0 lies in $x_1=0$, (31) shows that the corresponding surface in S_2 lies in $u=0$. Again, since $y_1\geqq 0$ on SS^0, it follows from (27) that $-[1-v^2]^{1/2}/2\leqq w\leqq[1-v^2]^{1/2}/2$; hence, SS^0 is represented by the ellipse E: $v^2+4w^2\leqq 1$. We have seen already that the boundary of E represents O_2^0.

An orbit on which $y_1>0$ corresponds to a stream line in S_2 on which u is increasing. When the orbit crosses the x_2-axis, the stream line in M^0 crosses SS^0, and the stream line in S_2 passes through E. On the other hand, if $y_1<0$ when the orbit crosses the x_2-axis, the stream line in M^0 does not cross SS^0, and the stream line in S_2 passes through $u=0$ on the exterior of E. It is thus possible to visualize the flow in M^0. We observe among other things that all stream lines cross the representation of SS^0 in S_2 in the same sense.

Let a transformation T be defined on SS^0 as follows: a stream line which crosses SS^0 at P has its next succeeding crossing at P' and its kth succeeding crossing at $P^{(k)}$. Then $P'=T(P)$ and $P^{(k)}=T^k(P)$. We proceed to study T.

It is possible to use (x_2, y_2) as coordinates on SS^0 since

$$(35)\qquad \begin{aligned} x_2 &= x_2, \qquad y_2 = y_2,\\ y_1 &= (1 - |\lambda_2|^2 x_2^2 - y_2^2)^{1/2} \end{aligned}$$

is merely a parametric representation of (33) with (x_2, y_2) as the parameters. Then T can be expressed in terms of (x_2, y_2).

Assume that the stream line has its first crossing at $t=0$; then from (30) the coordinates of P are

$$(36)\qquad x_2 = \frac{a_2}{|\lambda_2|}\sin\theta_2, \qquad y_2 = a_2\cos\theta_2,$$

and the coordinates of $P^{(k)}$ are

$$(37)\qquad \begin{aligned} x_2^{(k)} &= \frac{a_2}{|\lambda_2|}\sin\left(\left|\frac{\lambda_2}{\lambda_1}\right| 2k\pi + \theta_2\right),\\ y_2^{(k)} &= a_2\cos\left(\left|\frac{\lambda_2}{\lambda_1}\right| 2k\pi + \theta_2\right). \end{aligned}$$

From (36), (37) we see that T has the invariant function $F = |\lambda_2|^2 x_2^2 + y_2^2$, and that each of the curves $F = a_2^2$ is a path curve of T. From (35) it follows that this path curve is the ellipse $y_1 = [1 - a_2^2]^{1/2}$ on SS^0. By letting a_2 take on all values on $0 \leqq a_2 \leqq 1$, we get a family of ellipses which fill up SS^0. The path curves and F exist because the dynamical system is integrable [Birkhoff 3, pp. 114–115]. There are two integrals of (28) besides (29):

$$(38)\qquad |\lambda_i|^2 x_i^2 + y_i^2 = a_i^2 \qquad (i = 1, 2).$$

Of the three integrals, only two are independent.

In order to see more clearly the nature of T, we transform to new parameters (ξ, η) by means of

$$(39)\qquad \xi = |\lambda_2|\, x_2, \qquad \eta = y_2.$$

Corresponding to (36), (37) the coordinates of P, $P^{(k)}$ are now

$$(40)\qquad \xi = a_2 \sin\theta_2, \qquad \eta = a_2\cos\theta_2;$$

$$(41)\qquad \begin{aligned} \xi_k &= a_2\sin\left(\left|\frac{\lambda_2}{\lambda_1}\right| 2k\pi + \theta_2\right),\\ \eta_k &= a_2\cos\left(\left|\frac{\lambda_2}{\lambda_1}\right| 2k\pi + \theta_2\right). \end{aligned}$$

Expand the right hand member of (41) and substitute from (40). Then

$$
(42)\qquad
\begin{aligned}
\xi_k &= \quad \xi \cos\left(\left|\frac{\lambda_2}{\lambda_1}\right| 2k\pi\right) + \eta \sin\left(\left|\frac{\lambda_2}{\lambda_1}\right| 2k\pi\right),\\
\eta_k &= -\xi \sin\left(\left|\frac{\lambda_2}{\lambda_1}\right| 2k\pi\right) + \eta \cos\left(\left|\frac{\lambda_2}{\lambda_1}\right| 2k\pi\right),
\end{aligned}
$$

i.e., T^k is a rotation about P^0: $\xi=\eta=0$ of SS^0 into itself. The point P^0 is therefore invariant under T; (39) shows that it corresponds to O_1^0. If $|\lambda_2|/|\lambda_1|$ is a rational fraction p/q, p and q without common factors, then T^q rotates SS^0 through p complete revolutions, and every point is invariant. In this case, every trajectory is closed and periodic. If $|\lambda_2|/|\lambda_1|$ is irrational, P^0 is the only invariant point of T and its iterates, and the only closed and periodic orbits of the system are O_1^0, O_2^0. These results prove the following theorem.

THEOREM 13. *The transformation T on SS^0 is a rotation. The center of rotation P^0 is an invariant point which corresponds to O_1^0. If $|\lambda_2|/|\lambda_1|$ is rational, every trajectory of the system is closed and periodic; if it is irrational, only O_1^0 and O_2^0 are closed and periodic.*

Consider T on E in S_2. The path curves $y_1^2=a_1^2$ on SS^0 correspond to the curves

$$
(43)\qquad (1-v^2)^{1/2} \arc \tan\left[\pm\left(\frac{1-a_1^2-v^2}{a_1^2}\right)^{1/2}\right] = \pi w.
$$

As a_1 varies from 0 to 1, we get a family of simple closed curves beginning with the ellipse (32) and shrinking down to its center. The center of rotation on SS^0 corresponds to the center of E, which is therefore an invariant point. We have previously shown that the stream line representing O_1^0 crosses E at its center. We may thus picture T on E as a distorted rotation which carries each of the curves (43) into itself.

8. **A formula in geometry.** We turn aside from our main subject to prove a formula that was used in the last section.

A 2-dimensional surface in 4-space is defined by

$$
(44)\qquad f(x_1,\cdots,x_4)=0,\qquad \phi(x_1,\cdots,x_4)=0,
$$

and a unit vector $C:(c_i)$ has its initial end at the point (x_i^0) of the surface. The problem is to obtain a formula for the angle which C makes with (44).

DEFINITION. *The angle which C makes with the surface* (44) *is the complement of the angle between C and the normal to the surface which lies in the 3-plane containing C and the tangent 2-plane to the surface.*

We assume that the two 3-surfaces in (44) are not tangent at (x_i^0), i. e., we assume that the rank of the matrix

$$\left\| \begin{matrix} f_{x_i} \\ \phi_{x_i} \end{matrix} \right\| \tag{45}$$

is 2 [a subscript letter denotes a partial derivative with respect to that letter]. The tangent 2-plane to the surface at (x_i^0) is given by

$$f_{x_i}(x_i - x_i^0) = 0, \qquad \phi_{x_i}(x_i - x_i^0) = 0. \tag{46}$$

A repeated subscript in a product denotes a summation with respect to that subscript from 1 to 4. The two 3-planes in (46) are distinct since the rank of (45) is 2; taken together, therefore, they define a 2-plane.

Now determine the 3-plane which contains the tangent 2-plane (46) and the given vector C. All 3-planes which contain (46) are given by

$$Af_{x_i}(x_i - x_i^0) + B\phi_{x_i}(x_i - x_i^0) = 0. \tag{47}$$

This plane contains C if and only if it contains its end point $(x_i^0 + c_i)$. Substitute the coordinates of this point in (47) and solve for A, B; the result is

$$A = \phi_{x_i}c_i, \qquad B = -f_{x_i}c_i. \tag{48}$$

Now if A, B as given by (48) are both zero, we see that C lies in the tangent 2-plane (46) of the surface. Then C is tangent to the surface (44).

Assume henceforth that C does not lie in the tangent 2-plane; then A, B are not both zero and the required plane is

$$(\phi_{x_i}c_i)[f_{x_j}(x_j - x_j^0)] - (f_{x_i}c_i)[\phi_{x_j}(x_j - x_j^0)] = 0. \tag{49}$$

The ∞^1 normals to the surface at (x_i^0) have the direction components

$$\rho_1 f_{x_i} + \rho_2\phi_{x_i}. \tag{50}$$

Now determine ρ_1, ρ_2 so that (50) lies in (49). A point on the vector (50) is $(x_i^0 + \rho_1 f_{x_i} + \rho_2\phi_{x_i})$. Substitute in (49) and solve for ρ_1, ρ_2. The result is

$$\begin{aligned} \rho_1 &= \quad (f_{x_i}\phi_{x_i})(\phi_{x_j}c_j) - (\phi_{x_i}\phi_{x_i})(f_{x_j}c_j), \\ \rho_2 &= -(f_{x_i}f_{x_i})(\phi_{x_j}c_j) + (\phi_{x_i}f_{x_i})(f_{x_j}c_j). \end{aligned} \tag{51}$$

Substitute these values for ρ_1, ρ_2 in (50), and we have the required normal vector $N:(n_i)$. Then if ψ is the angle at which C crosses the surface (44),

$$\cos\left(\frac{\pi}{2} - \psi\right) = \frac{(c_i n_i)}{(n_i n_i)^{1/2}}. \tag{52}$$

One detail remains. It must be shown that ρ_1, ρ_2 are not both zero, for if

they were, N would be a null vector. The desired result follows from (51) when we assume that (45) is of rank 2, and that A, B in (48) are not both zero.

9. **Some results on surfaces of section and surface transformations.** A common type of surface of section for reversible dynamical systems with two degrees of freedom is formed as follows: Take a closed orbit O without multiple points which either has no point in common with Z, or is an orbit traced with a backward and forward motion between two points of Z. Consider all points $(x_1, x_2; y_1, y_2)$ in M such that (x_1, x_2) is a point of O and (y_1, y_2) is a velocity vector which is tangent to O or lies on a certain specified side of it. These points form a surface Σ.

THEOREM 14. *The surface Σ is an analytic surface.*

Suppose first that O does not pass through a point of Z. Take any point P on O and rotate the axes so that the tangent at P is parallel to the x_2-axis. Then the equation of O near P can be written in the form $x_1=\phi(x_2)$, where ϕ is analytic. Suppose Σ is formed with the velocity vectors for which $y_1 \geqq y_2\phi'(x_2)$, the prime here denoting a derivative with respect to x_2. Then Σ is defined by

$$(53)\qquad \begin{gathered} M(x_1, x_2; y_1, y_2) \equiv T - U - h/2 = 0, \\ x_1 = \phi(x_2), \qquad y_1 \geqq y_2\phi'(x_2). \end{gathered}$$

We propose to show that one or the other of the sets (x_2, y_1), (x_2, y_2) can be taken as the parameters of an analytic representation of Σ in the neighborhood of P. Substitute from the second equation in (53) in the first. Then if the equation $M[\phi(x_2), x_2; y_1, y_2]=0$ gives either y_1 or y_2 as an analytic function of the other two variables, the desired result follows. Now $M=0$ can be solved for y_i if $\partial M/\partial y_i \equiv \partial T/\partial y_i \neq 0$. The desired result follows then unless both of these partial derivatives vanish. But $\partial T/\partial y_1$, $\partial T/\partial y_2$ vanish simultaneously only at points on Z, and O has no point in common with Z by hypothesis. Hence, Σ is analytic in the neighborhood of P, and since P was any point of O, Σ is analytic throughout.

Now suppose that O joins two points of Z. The proof given above applies to any interior point of O; hence, it will be sufficient to show that the part of Σ arising from points of O near Z is analytic. The orbit is given to us from the equations of motion in the parametric form

$$(54)\qquad x_1 = x_1(t), \qquad x_2 = x_2(t).$$

If O touches Z at $P^0:(x_1^0, x_2^0)$, it is not regular there, i.e., both x_1' and x_2' vanish there. We shall show, however, that this state of affairs results from

the fact that the particle reverses its direction of motion there, and not from the nature of the curve itself.

Now if O passes through P^0 at time $t=0$, equations (54) are

$$x_1 = x_1^0 + a_2t^2 + a_4t^4 + \cdots,$$
$$x_2 = x_2^0 + b_2t^2 + b_4t^4 + \cdots, \tag{55}$$

only even powers of t occurring, because the equations of motion and the initial conditions

$$x_1 = x_1^0, \; y_1 = 0, \; x_2 = x_2^0, \; y_2 = 0, \qquad t = 0,$$

are unchanged when t is replaced by $-t$. Let T and U be

$$T = \frac{1}{2}\sum_{i,j=1}^{2} T_{ij}(x_1, x_2)x_i' x_j', \qquad T_{ij} = T_{ji},$$
$$U = U(x_1, x_2).$$

Since T is a positive definite quadratic form, we have

$$| T_{ij} | > 0. \tag{56}$$

From the equations of motion, we find that a_2, b_2 satisfy the equations

$$T_{11}(x_1^0, x_2^0)a_2 + T_{12}(x_1^0, x_2^0)b_2 = U_{x_1}(x_1^0, x_2^0),$$
$$T_{21}(x_1^0, x_2^0)a_2 + T_{22}(x_1^0, x_2^0)b_2 = U_{x_2}(x_1^0, x_2^0). \tag{57}$$

But since there are no double points on Z by hypothesis [see (8)], we see that a_2, b_2 are not both zero. Suppose $b_2 \neq 0$. Then it is possible to solve the second equation in (55) for t^2, obtaining an analytic function of $(x_2 - x_2^0)$. Use this function to eliminate t^2 from the first equation in (55). We obtain

$$x_1 = \phi(x_2 - x_2^0) \equiv x_1^0 + \frac{a_2}{b_2}(x_2 - x_2^0) + \cdots, \tag{58}$$

which defines a real analytic curve which crosses Z. The orbit is formed from the part of this curve which lies in R. The irregularity at P^0 is therefore due to the reversal of the direction of motion and not to the nature of the curve itself.

Furthermore, the curve (58) is not tangent to Z at P^0. Using the values of a_2, b_2 as given by (57), we find that the curves are tangent if and only if

$$T_{22}U_{x_1}^2 - 2T_{12}U_{x_1}U_{x_2} + T_{11}U_{x_2}^2 = 0. \tag{59}$$

But this is impossible, because there are no double points on Z, and the quadratic form is positive definite. Also, (56) and (57) show that if we choose

the axes so that $U_{x_1}=0$ at P^0, then $b_2\neq 0$, and the equation of the curve along which O lies can be written in the form (58) near P^0.

We can now complete the proof that Σ is analytic. Rotate the axes so that Z is parallel to the x_1-axis at P^0. Then $U_{x_1}=0$ at P^0, and near this point the equation of the curve along which O lies can be written in the form (58). Furthermore Σ is defined by the equations

$$\begin{gathered} M(x_1, x_2; y_1, y_2) = 0, \\ x_1 = \phi(x_2 - x_2^0), \qquad y_1 \geqq y_2\phi'(x_2 - x_2^0). \end{gathered} \tag{60}$$

Then (y_1, y_2) can be taken as the parameters on Σ, for substitute from the second equation in (60) in the first. The resulting equation can be solved for x_2 if its partial derivative with respect to x_2 is not zero. At P^0 this partial derivative reduces to $U_{x_2}\neq 0$; hence, x_2 can be expressed analytically in terms of (y_1, y_2). Substitute now for x_2 in the second equation in (60), and we have x_1 also expressed analytically in terms of (y_1, y_2). Thus we have proved that Σ is analytic in all cases.

THEOREM 15. *The angle at which a trajectory crosses Σ is of the first order in the distance to the boundary.*

Now it can be shown that $(y_1-y_2\phi')$ is an infinitesimal of the first order in the distance from a point on Σ to the boundary. Also, Σ is defined by equations and inequalities of which (53) are typical. The direction components of a stream line are dx_1/dt, dx_2/dt, dy_1/dt, dy_2/dt. Let ψ be the angle which the stream line makes with Σ at the point of crossing. Now the two 3-dimensional surfaces in (53) and (60) are never tangent since O is never tangent to Z. Then the formula developed in §8 can be used for determining ψ. Remembering that $dM/dt\equiv 0$ because $M=0$ is the integral of energy, we find by a straightforward calculation that

$$\sin\psi = \frac{-(M_{x_1}^2 + M_{x_2}^2 + M_{y_1}^2 + M_{y_2}^2)(y_1 - y_2\phi')}{\left[\left(\frac{dx_1}{dt}\right)^2+\left(\frac{dx_2}{dt}\right)^2+\left(\frac{dy_1}{dt}\right)^2+\left(\frac{dy_2}{dt}\right)^2\right]^{1/2}(m_im_i)^{1/2}}. \tag{61}$$

Here $(m_1, \cdots, m_4)$ denotes a vector which differs from the vector $(n_1, \cdots, n_4)$ of §8 only by a factor. An examination shows that it is not a null vector, not even on the boundary of Σ. Since there are no double points on Z, the first factor in the numerator is not zero. For the same reason, the stream lines are regular curves at every point, and the first radical in the denominator does not vanish. Then the theorem follows immediately from the fact that $(y_1-y_2\phi')$ measures the distance to the boundary of Σ.

If O does not touch Z, Σ is bounded by two closed stream lines corresponding to O traced in the two directions, and it is homeomorphic to a ring bounded by two concentric circles. If O joins two points of Z, Σ is bounded by a single closed stream line which corresponds to O, and it is homeomorphic to a circular disc. In all cases Σ is an analytic surface; the angle at which a stream line crosses Σ is of the first order in the distance to the boundary; and all stream lines which cross Σ cross it in the same sense. Then if it can be shown that every trajectory of the system crosses Σ at least once in a given interval of time, it follows from the definition that Σ is a surface of section.

We proceed to the proof of a theorem which gives an important qualitative result on the nature of the orbits of a reversible dynamical system.

THEOREM 16. *If there exists a periodic orbit O joining two points of Z from which a surface of section SS of type Σ can be formed, and if R is homeomorphic to a circular disc, there exists at least one further periodic orbit joining two points of Z.*

In the first place, since O joins two points of Z, SS is bounded by a single closed stream line and is homeomorphic to a circular disc. Since SS is a surface of section, there is an analytic transformation T on it defined in the usual way. Furthermore, T has a certain number of invariant points [Birkhoff 1, p. 287]. To each invariant point P_i an integer δ_i is assigned as follows: Draw the vector from a given point Q in the neighborhood of P_i to its image Q' under T. Then when Q describes a small circle about P_i in the positive direction, the vector rotates through the angle $2\delta_i\pi$. Now the sum of the δ_i for all the invariant points on SS is 1 [Birkhoff 1, p. 290; note that the formula should be $(2q+d-2)$].

Assume now that the theorem is false, i.e., assume that O is the only periodic orbit which joins two points of Z. Then each invariant point of T arises from a closed orbit which does not touch Z. Corresponding to such orbits there are two stream lines in M and they are distinct [Theorem 2]. Since O divides R into two regions, each of these stream lines crosses SS and gives rise to the same number of invariant points of T. Hence, to each invariant point P_i there is a unique second invariant point Q_i; the number of invariant points is even.

Now if it can be shown that P_i and Q_i have the same number δ_i, it will follow that the sum of the δ_i is an even number in contradiction to the fact that it is 1. But δ_i is determined from the equations of variation and is the same for P_i and Q_i. The theorem follows.

At the same time we have proved that under the hypotheses of Theorem 16, the following theorem is true also.

THEOREM 17. *The invariant points of T and its powers are paired, the two points of a pair being distinct unless the corresponding orbit joins two points of Z.*

10. Analytic continuation of the surface of section. For the present assume only that $\lambda_1 \neq \lambda_2$. Then by Corollary 1, §5, at least one of the two orbits O_1^0, O_2^0 can be continued analytically for $\mu > 0$. Suppose the notation is so chosen that it is O_2^0 which can be continued. Then for $\mu = 0$ the system has the surface of section SS^0 as described in §7.

For $\mu = 0$ the periodic orbit O_2 reduces to O_2^0, which lies along the line $x_1 = 0$; since O_2 varies analytically with μ, it follows that for μ sufficiently small its equation can be written in the form

$$x_1 = \phi(x_2, \mu), \qquad \phi(x_2, 0) \equiv 0. \tag{62}$$

We shall now show that for each value of μ the surface in M defined by (62) and the inequality $y_1 \geqq y_2 \partial\phi/\partial x_2$ forms a surface of section SS which is the analytic continuation of SS^0.

The surface SS is a surface of the type Σ studied in §9. It is an analytic surface in M which varies analytically with μ and reduces to SS^0 for $\mu = 0$. As shown in §9, it forms a surface of section if every trajectory cuts it at least once in a fixed interval θ of time. It was shown in §7 that this requirement is satisfied for $\mu = 0$. Now the intersections of a given stream line with SS are determined by the intersections of the corresponding orbit with (62). These orbits intersect at an angle different from zero for $\mu = 0$, and since they vary analytically with μ, they continue to intersect at least for $\mu > 0$ but small, with the length of time between successive crossings of SS uniformly bounded. Conceivably this argument might fail in the neighborhood of the boundary of SS, but here we have recourse to the equations of variation for the trajectory corresponding to (62). A detailed consideration shows that every trajectory continues to cross SS for μ sufficiently small, and that the length of time between successive crossings is uniformly bounded; hence, SS is a surface of section as stated.

The definition of T on SS is the same as in previous cases. Now R is homeomorphic to a circular disc [§6], and all the other hypotheses of Theorem 16 are satisfied. This theorem proves that the equations corresponding to (24) have a solution for each value of μ so long as the surface of section exists. Continuation of O_1^0 is therefore possible, but our results do not show that this continuation is unique.

THEOREM 18. *If $\lambda_1 \neq \lambda_2$, there exist at least two periodic orbits joining two points of Z for $\mu > 0$ but small.*

Suppose now that neither λ_1/λ_2 nor λ_2/λ_1 is an integer. Then by Theorem 8 both O_1^0 and O_2^0 can be continued analytically. There is an invariant point P on SS which corresponds to O_1, the continuation of O_1^0. For $\mu=0$, T is a rotation about P^0 through an angle not an integral multiple of 2π. Then δ for P^0 is 1, and P^0 is said to be a simple and stable invariant point [Birkhoff 1, pp. 287–288]. Now the nature of T about P is determined by the characteristic exponents for the corresponding trajectory [Poincaré 1, vol. I, chap. IV]. The characteristic exponents are continuous functions of μ; hence, for $\mu>0$ but small, P is a stable invariant point. Then it is possible to consider SS a ring surface with P the inner boundary. We shall show that under certain conditions Poincaré's Last Geometric Theorem can be applied to the transformation on this ring [Poincaré 2; Birkhoff 2, and 1, p. 294].

In the first place, T is an analytic transformation of the ring into itself which carries the boundaries into themselves. In the second place, T has an invariant area integral as a result of Theorem 11 [Birkhoff 1, p. 285]. Finally, there are rotation numbers α_B and α_P associated with the transformation on the two boundaries of the ring [Birkhoff 3, pp. 87–88]. We have shown in §7 that for $\mu=0$ the two rotation numbers are equal. We shall show by means of an example, however, that α_B and α_P are functions of μ and in general are not equal. Then Poincaré's considerations prove the following theorem [Poincaré 2, §3; Birkhoff 1, pp. 297–298].

THEOREM 19. *If $\alpha_B\neq\alpha_P$, there exist infinitely many periodic orbits.*

The limiting integrable system affords a good example to show that the conclusion of this theorem may not hold if $\alpha_B=\alpha_P$ [see §7].

As an example to show that α_B and α_P are not identically equal, consider the system for which the equations of motion and the integral of energy are

$$
\begin{aligned}
x_1'' &= -|\lambda_1|^2x_1 + 2\mu^2x_1^3,\\
x_2'' &= -|\lambda_2|^2x_2,\\
y_1^2 + y_2^2 &+ |\lambda_1|^2x_1^2 + |\lambda_2|^2x_2^2 - \mu^2x_1^4 = 1.
\end{aligned}
\tag{63}
$$

For all values of μ there are two periodic orbits O_1, O_2 joining two points of Z, and they lie along the two axes. The surface of section SS is defined by $x_1=0$, $y_1\geqq 0$. As in the limiting system $\mu=0$, we may take (x_2, y_2) as the coordinates on the surface.

For all values of μ the point $x_2=y_2=0$ is the invariant point P on SS and corresponds to O_1. The equation of O_1 is $x_1=x_1(t, \mu)$, $x_2=0$; its period is $[2\pi/|\lambda_1| + \tau_1(\mu)]$, where $\tau_1(0)=0$. The period is independent of the amplitude only for simple harmonic motion, however; hence, $\tau_1(\mu)\neq 0$.

Consider α_P first. A nearby orbit to O_1 is $x_1 = x_1(t, \mu) + \xi_1$, $x_2 = \xi_2$, and from (63) we find that the equations of variation are

$$\text{(64)} \qquad \begin{aligned} \xi_1'' &= -|\lambda_1|^2\xi_1 + 6\mu^2[x_1(t, \mu)]^2\xi_1, \\ \xi_2'' &= -|\lambda_2|^2\xi_2. \end{aligned}$$

From the second of these equations we obtain

$$\text{(65)} \qquad \begin{aligned} \xi_2 &= \xi_2^0 \cos |\lambda_2| t + (\eta_2^0/|\lambda_2|) \sin |\lambda_2| t, \\ \eta_2 &= -|\lambda_2| \xi_2^0 \sin |\lambda_2| t + \eta_2^0 \cos |\lambda_2| t. \end{aligned}$$

Here (ξ_2, η_2) are coordinates in the neighborhood of P and correspond to (x_2, y_2), and (ξ_2^0, η_2^0) is the point on SS through which the stream line passes at time $t=0$. To find the point into which it is carried by T, we have only to set $t = [2\pi/|\lambda_1| + \tau_1(\mu)]$ in (65). By introducing new coordinates as was done in §7, we show that the limiting transformation at P is a rotation through the angle

$$\text{(66)} \qquad -|\lambda_2| [2\pi/|\lambda_1| + \tau_1(\mu)].$$

Then α_P is given by (66).

Now consider α_B. The orbit corresponding to the stream line which forms the boundary B is $x_1 = 0$, $x_2 = \sin |\lambda_2| t$, which is periodic with period $2\pi/|\lambda_2|$. The equations of variation are found in the usual way to be

$$\text{(67)} \qquad \begin{aligned} \xi_1'' &= -|\lambda_1|^2\xi_1, \\ \xi_2'' &= -|\lambda_2|^2\xi_2. \end{aligned}$$

The first equation determines the intersections of the varied stream line with SS; the stream line crosses SS when $\xi_1 = 0$, $\eta_1 \geqq 0$. As in previous cases, we can use (ξ_2, η_2) as coordinates on SS near B. Then equations (65) hold. A stream line which crosses SS at $t=0$ has its next crossing at $t = 2\pi/|\lambda_1|$. Then (65) show that T on B is essentially a rotation through the angle

$$\text{(68)} \qquad -2\pi|\lambda_2|/|\lambda_1|.$$

Then α_B is given by (68).

Comparing (66) and (68), we see that α_P and α_B are in general not equal since $\tau_1(\mu) \neq 0$. Our conclusion is the following: *If the system really depends on μ, i.e., if it is not identical with the limiting integrable system for all values of μ, then α_P and α_B are not equal in general for $\mu > 0$.*

11. **Symmetric systems.** We shall now suppose that the system is symmetric in the origin on the characteristic surface, i.e., we assume that (9) holds. As we have already seen, a system of this kind has special properties, which we shall now study in greater detail.

Assume now that neither $|\lambda_1|/|\lambda_2|$ nor $|\lambda_2|/|\lambda_1|$ is an odd integer. Then by Theorem 9 both O_1^0 and O_2^0 can be continued analytically for $\mu>0$; for each value of μ these orbits pass through the origin and are symmetric in this point. A surface of section SS can be formed from either one of these orbits; let it be formed from O_2. All the results of §10 apply in the present case, but the symmetry leads to special properties of T. In order to state the results more easily, we employ the representation of M in S_2 [see §6].

It is clear that R can be deformed into $C: u^2+v^2\leqq 1$ in such a way that symmetric points are carried into points symmetric in the center of C, and so that O_1 and O_2 lie along the diameters $v=0$ and $u=0$ of C respectively. Then the stream lines in S_2 have the essential properties of symmetry possessed by the stream lines in M; also SS is represented in S_2 by the ellipse E in the plane $u=0$.

Now it was shown in §2 that the stream lines are related by fours. The significance of this fact is that there are two transformations V_1 and V_2 which when applied to a stream line and its transforms by V_1 and V_2 yield four and only four stream lines. In S_2 the transformation V_1 is

$$
\begin{aligned}
u' &= u, \qquad v' = v, \\
w' &= w + [1 - (u^2 + v^2)]^{1/2},
\end{aligned}
\tag{69}
$$

coupled with a reversal of the direction of flow [see Theorem 1 and (27)]. By reversal of the direction of flow, we mean the following: if the flow proceeds from P to Q on the given stream line T^*, it proceeds from Q' to P' on V_1T^*. If necessary w' in (69) is to be reduced modulo $2[1-(u^2+v^2)]^{1/2}$. As shown in §2, T^* and V_1T^* correspond to a single orbit traced in the two directions. Obviously $V_1^2=I$, the identity.

In S_2 we find that V_2 is defined by

$$
\begin{aligned}
u' &= -u, \qquad v' = -v, \\
w' &= w + [1 - (u^2 + v^2)]^{1/2},
\end{aligned}
\tag{70}
$$

without reversal of the direction of flow [see Theorem 4 and (27)]. Again w' is to be reduced modulo $2[1-(u^2+v^2)]^{1/2}$ when necessary. Then $V_2^2=I$.

Now V_1 and V_2 generate a group with the four distinct transformations I, V_1, V_2, V_1V_2. We see that V_1V_2 is merely a reflection in the w-axis in S_2 with reversal of the direction of flow, and that $(V_1V_2)^2=I$. By applying the transformations of this group to T^*, we obtain three others. The four stream lines are permuted among themselves by any transformation of the group. Now the four are not always distinct [see §2]. If the orbit corresponding to any one of them is its own symmetric image or is a curve touching Z, there are

at most two distinct stream lines; if it is both, the four are identical. In all other cases the four are distinct.

Now if T^* is closed, the other three stream lines of a group are closed; hence, with the obvious convention in case they are not all distinct, we have the following theorem.

THEOREM 20. *The invariant points of T and its powers and the closed periodic trajectories of the system occur in groups of four.*

Now SS is represented by E in the plane $u=0$; hence, T may be expressed in terms of the coordinates (v, w). Since (V_1V_2) carries a stream line into a stream line with reversal of the direction of flow, we see that *if T^k carries (v_0, w_0) into (v_1, w_1), then T^k also carries $(-v_1, w_1)$ into $(-v_0, w_0)$*. It follows that $(V_1V_2)T(V_1V_2)T=I$. Then $(V_1V_2)T$ is a transformation U with period 2: $(V_1V_2)T=U$, $U^2=I$. Hence, $T=(V_1V_2)U$. We have thus proved the following theorem.

THEOREM 21. *The transformation T is the product of two transformations, one of which is a reflection in the w-axis, and both of which have the period* 2.

Now suppose that T^k carries a point $(0, w_0)$ into $(0, w_1)$; then by the italicized statement above, T^k also carries $(0, w_1)$ into $(0, w_0)$, and both points are invariant under T^{2k}. They correspond to a single closed stream line. Let the segment $v=0$ on E be denoted by AB. A point on AB corresponds to an orbit passing through the center of symmetry, and the above statement is equivalent to Theorem 5. Thus, if we can prove that there are points on AB which are transformed into points on AB, we can conclude that there exist closed periodic orbits passing through the center of symmetry.

In the first place, there exists an invariant point P of T on AB. It is the center of E, the point at which the stream line corresponding to O_1 crosses E. Now consider the images of AB on E under T and its powers. If the rotation numbers α_P and α_B [see §10] are unequal, the image of AB under T and its powers is a spiral, and we can assert that there is an integer N such that the image of AB under T^k for $k \geqq N$ intersects AB in points distinct from P. This is proved as follows. The rotation numbers for T^k on the boundary B of E and at P are $k\alpha_B$ and $k\alpha_P$. Then for k sufficiently large, say $k=N$, $k\alpha_B$ and $k\alpha_P$ correspond to transformations differing by at least one complete cycle; for $k=\rho N$, they differ by at least ρ cycles. Hence, the image of AB under T^k intersects AB at least for $k \geqq N$, and the number of such intersections becomes infinite with k.

Suppose that Q_1 on AB is carried into Q_2, distinct from Q_1, on AB by T^m, and that m is the smallest power of T for which this happens. Then Q_1, Q_2

are invariant under T^{2m} and correspond to a periodic orbit O which passes twice through the center of symmetry; the two branches there have distinct tangents since Q_1, Q_2 are distinct. The corresponding stream line crosses SS $2m$ times. Since R is divided into two parts by O_2, we see that O crosses O_2 twice for each crossing of SS by the stream line, i.e., O crosses O_2 $4m$ times. We therefore say this orbit is of type O_{4m}. An orbit O_{4m} cannot touch Z, because it has two distinct branches at the center of symmetry. We have thus proved the following theorem.

THEOREM 22. *If $\alpha_B \neq \alpha_P$, there exists an infinite number of closed periodic orbits of type O_{4m}, there being one or more for each $m \geqq N$.*

We proceed to establish the existence of an infinite number of periodic orbits of a second type.

The transformation

$$x_i' = x_i, \qquad y_i' = -y_i \qquad (i = 1, 2), \tag{71}$$

transforms M into itself and in particular carries SS into a surface SS' which is also a surface of section. The surfaces SS and SS' are bounded by the same closed stream line, and taken together they form a surface homeomorphic to a 2-sphere. In S_2 the transformation (71) is V_1; hence, SS' is represented in S_2 by E', the part of the circle $u=0$ which lies outside of E. By means of (71) we extend the definition of T on SS to the combined surface $SS+SS'$. Each half of this surface is a surface of section; hence, the stream lines define a transformation of it into itself. Since O_2 divides R into two parts, a stream line which crosses SS (SS') at Q has its first succeeding crossing of the surface at Q' on SS' (SS). Then the new transformation is that which carries Q into Q'; we designate it by $T^{1/2}$ since it has the obvious property that its square is T. Also $T^{1/2}$ is not a sense-preserving transformation; it has the nature of a reflection.

We return to the representation in S_2 in order to simplify the exposition. The stream line which crosses E at P also crosses E' at P'. Then $P'=T^{1/2}(P)$. Associated with the transformation of P into P' by T there is a rotation number which is obviously $\alpha_P/2$; the common boundary B of E and E' is transformed into itself by $T^{1/2}$, and the corresponding rotation number is $\alpha_B/2$.

Suppose $T^{k+1/2}$, k an integer, carries (v_0, w_0) into (v_1, w_1). Apply the transformation (V_1V_2), and we see that $T^{k+1/2}$ also carries $(-v_1, w_1)$ into $(-v_0, w_0)$. Then as before $(V_1V_2)T^{1/2}(V_1V_2)T^{1/2}=I$. Set $(V_1V_2)T^{1/2}=V$. Then $T^{1/2}=(V_1V_2)V$, where $V^2=I$. Thus by defining T on the entire surface $SS+SS'$

we are able to factor T into the product of four factors, each with the period $2: T=(V_1V_2)V(V_1V_2)V$.

Now if $T^{k+1/2}$ carries $(0, w_0)$ into $(0, w_1)$, it also carries $(0, w_1)$ into $(0, w_0)$, and both points are invariant under T^{2k+1}. One of these points is on E. Let $A'B'$ designate the line segment $v=0$ on E'. Then if we can show that the image of AB under $T^{k+1/2}$ intersects $A'B'$, we can prove the existence of further periodic orbits. Now the image of AB under $T^{k+1/2}$ is a spiral which always intersects $A'B'$ at P'. For k sufficiently large, and at least for $k>N$, this spiral intersects $A'B'$ at points other than P'. The number of such intersections becomes infinite with k. Each such intersection gives an invariant point on E under T^{2k+1}.

Suppose that Q is such an invariant point on E under T^{2m+1}, and that this is the lowest power of T under which it is invariant. The corresponding orbit O passes twice through the center of symmetry and crosses O_2 $(4m+2)$ times. This orbit is therefore said to be of type O_{4m+2}. The orbits of this type may or may not touch Z. The orbit O_1, which corresponds to P on E, is included in this class with $m=0$. We have thus proved the following theorem.

THEOREM 23. *If $\alpha_B \neq \alpha_P$, there exist an infinite number of periodic orbits of type O_{4m+2}, there being one or more for each $m \geqq N$.*

Similar results can be obtained if it is assumed that the system is symmetric in one or both of the axes on the characteristic surface.

Bibliography

BIRKHOFF, G. D.

1. *Dynamical systems with two degrees of freedom*, these Transactions, vol. 18 (1917), pp. 199–300.
2. *Proof of Poincaré's geometric theorem*, these Transactions, vol. 14 (1913), pp. 14–22.
3. *Surface transformations and their dynamical applications*, Acta Mathematica, vol. 43 (1922), pp. 1–119.
4. *Dynamical Systems*, American Mathematical Society Colloquium Publications, vol. 9 (1927).
5. *The restricted problem of three bodies*, Rendiconti del Circolo Matematico di Palermo, vol. 39 (1915), pp. 265–334.
6. *On the periodic motions of dynamical systems*, Acta Mathematica, vol. 50 (1927), pp. 359–379.

HORN, J.

1. *Beiträge zur Theorie der kleinen Schwingungen*, Zeitschrift für Mathematik und Physik, vol. 48 (1902–03), pp. 400–434.

MORSE, M.

1. *The foundations of a theory in the calculus of variations in the large*, these Transactions, vol. 30 (1928), pp. 213–274.
2. *On a one-to-one representation of geodesics on a surface of negative curvature*, American Journal of Mathematics, vol. 43 (1921), pp. 33–51.
3. *Recurrent geodesics on a surface of negative curvature*, these Transactions, vol. 22 (1921), pp. 84–100.

MOULTON, F. R.

1. *Differential Equations*, Macmillan, 1930.

Painlevé, P.
1. *Sur les petits mouvements périodiques des systèmes*, Comptes Rendus, vol. 124 (1897), pp. 1222–1225.
Poincaré, H.
1. *Méthodes Nouvelles de la Mécanique Céleste*, vol. I (1892), vol. III (1899), Gauthier-Villars.
2. *Sur un théorème de géométrie*, Rendiconti del Circolo Matematico di Palermo, vol. 33 (1912), pp. 375–407.
Price, G. B.
1. *A class of dynamical systems on surfaces of revolution*, American Journal of Mathematics, vol. 54 (1932), pp. 753–768.
2. *On the Strömgren-Wintner natural termination principle*, American Journal of Mathematics, vol. 55 (1933), pp. 303-308.
Whittaker, E. T.
1. *Analytical Dynamics*, 3d edition (1927), Cambridge University Press.

University of Rochester,
Rochester, N. Y.

LAPLACE INTEGRALS AND FACTORIAL SERIES IN THE THEORY OF LINEAR DIFFERENTIAL AND LINEAR DIFFERENCE EQUATIONS*

BY

W. J. TRJITZINSKY

1. **Introduction.** Our present object is to carry out application of Laplace integrals (leading to convergent factorial series developments) to the fullest possible extent in the field of linear differential equations,

$$\text{(A)} \qquad L(y) \equiv \sum_{k=0}^{n} d_{n-k}(x) y^{(k)}(x) = 0;$$

and, secondly, in the field of linear difference equations

$$\text{(B)} \qquad L(y) \equiv \sum_{k=0}^{n} d_{n-k}(x) y(x + k) = 0.$$

In both (A) and (B) the coefficients are given by convergent series of the form

$$(1) \qquad d_{n-k}(x) = \sum_{s=-m}^{\infty} d_{n-k,s} x^{-s/p} \qquad (\text{integer } p \geqq 1),$$

and $d_0(x)$, $d_n(x) \not\equiv 0$. In view of the fact that the two mentioned fields are to a considerable degree analogous, it has been found convenient to give the developments for both of these in a single paper.

Another outstanding method, a method which previously had been applied with complete success in the two indicated fields, as well as in the field of q-difference equations, is that based on the study of asymptotic properties of solutions. In this connection the starting point is given by the full sets of formal series solutions (in general divergent), which are known to exist in all cases. Existence of a complete set of formal series solutions for a difference equation (B), as well as other results of a formal character, have been established by G. D. Birkhoff.† For differential equations (A) existence of a full set of formal solutions follows from a work of E. Fabry.‡ In the

* Presented to the Society, December 27, 1934; received by the editors April 18, 1934.

† *Formal theory of irregular linear difference equations*, Acta Mathematica, vol. 54 (1930), pp. 205–246.

‡ *Sur les intégrales des équations différentielles linéaires à coefficients rationnels*, Thèse, 1885, Paris.

asymptotic (analytic) theories of linear differential, difference and q-difference equations, respectively, the coefficients in the involved equations are either representable by convergent series of the form (1) or, more generally, they are merely asymptotic (in certain regions) to such, possibly divergent, series. In all essential particulars these three theories have been completely treated (*under no restrictions on the formal series solutions*) as follows. A joint work by G. D. Birkhoff and W. J. Trjitzinsky gives the developments for difference equations (B).* A paper by Trjitzinsky establishes the general asymptotic theory for q-difference equations.† Finally, the general asymptotic theory for differential equations (A) has been developed by Trjitzinsky.‡ Broadly speaking, at the basis of the several general asymptotic theories, referred to above, to a very substantial degree lie the ideas and methods due to Birkhoff or inspired by his work.

Important earlier developments, especially in the field of difference equations, are due to C. R. Adams, R. D. Carmichael, H. Galbrun, E. Hilb, O. Perron, S. Pincherle, H. Späth§ and some others.

While the asymptotic theory has been shown to yield full sets of analytic solutions (with appropriate asymptotic properties) in all cases, the situation is different inasmuch as the method of Laplace integrals and application of

---

* *Analytic theory of singular difference equations*, Acta Mathematica, vol. 60 (1932), pp. 1–89.

† *Analytic theory of linear q-difference equations*, Acta Mathematica, vol. 61 (1933), pp. 1–38; also cf. *The general case of integro-q-difference equations*, Proceedings of the National Academy of Sciences, vol. 18 (1932), pp. 713–719.

‡ *Analytic theory of linear differential equations*, Acta Mathematica, vol. 62 (1934), pp. 167–226.

§ R. D. Carmichael, for instance, chapter IV of the book referred to at the end of the next footnote. C. R. Adams, *On the irregular cases of the linear ordinary difference equation*, these Transactions, vol. 30 (1928), pp. 507–554. H. Galbrun, *Sur certains solutions exceptionnelles d'une équation linéaire aux différences finies*, Bulletin de la Société Mathématique de France, vol. 49 (1921), pp. 206–241. E. Hilb, *Zur Theorie der linearen Differenzengleichungen*, Mathematische Annalen, vol. 85 (1922), pp. 89–98; same title 1, Mathematische Zeitschrift, vol. 14 (1922); same title 2, ibid., vol. 15 (1922), pp. 280–285; same title 3, ibid., vol. 19 (1924), pp. 136–144. O. Perron, *Über lineare Differenzengleichungen zweiter Ordnung* . . . , Heidelberger Sitzungsberichte (mathematisch-physikalische Klasse), No. 17 (1917); *Über das Verhalten der Integrale linearer Differenzengleichungen im Unendlichen*, Jahresbericht der Deutschen Mathematiker-Vereinigung, vol. 19 (1910), pp. 129–137; *Über lineare Differenzengleichungen*, Acta Mathematica, vol. 34 (1910), pp. 109–137; *Über Systeme* . . . , Journal für die reine und angewandte Mathematik, vol. 147 (1917), pp. 36–53; *Über Summengleichungen und Poincarésche Differenzengleichungen*, Mathematische Annalen, vol. 84 (1921), pp. 1–15. S. Pincherle, *Sopra una transformazione delle equazioni differenziali lineari* . . . , Rendiconti, Istituto Lombardo, (2), vol. 19 (1886), pp. 559–562; *Sur la génération de systèmes récurrents* . . . , Acta Mathematica, vol. 16 (1892), pp. 341–363; *Sulla risoluzione dell' equazione funzionale* . . . , Memorie, Bologna Accademia, (4), vol. 9 (1888), pp. 181–204; *Sulle equazioni alle differenze*, Lincei Rendiconti, 1894, pp. 12–17, pp. 99–105; *Sulla risoluzione approssimata delle equazioni alle differenze*, Lincei Rendiconti, 1898, pp. 230–234. H. Späth, *Über das asymptotische Verhalten der Lösungen nichthomogener linearer Differenzengleichungen*, Acta Mathematica, vol. 51 (1927), pp. 133–199; same title, Mathematische Zeitschrift, vol. 30 (1929), pp. 487–513.

factorial series is concerned. This method, whenever successful, enables one to express a formal series solution with the aid of convergent factorial series. In every such case we have a situation when a possibly divergent formal series solution is "summed" by an essentially "exponential" method; however, this method is known to be applicable not in all cases. Basically and predominantly developments of this type rest on numerous important works of N.E. Nörlund* in the theory of factorial series and in connection with application of these series to difference equations. Among the works of others, involving application of Laplace integrals and factorial series to differential and difference equations, outstanding is a sequence of papers due to J. Horn.†

As will be seen from the Main Theorems (§§7, 12) the program of applying Laplace integrals and factorial series to equations (A) and (B) is capable of being extended considerably beyond the results of the earlier writers. On the other hand, certain examples (in §§7 and 12) will demonstrate the fact that these theorems cannot be extended (in a certain sense).

It is also to be noted that, in view of the reciprocal relationship between equations (A) and (B) on one side and corresponding linear systems on the other, results of the type established in these pages will hold for linear systems as well.

Numerous works, which we have not mentioned explicitly, are referred to in the several papers and books indicated in the footnotes of this introduction. However, we have indicated directly the more relevant ones of the previous contributions.

Part I. Linear differential equations

2. Some preliminary facts concerning differential equations. An equation (A) possesses a full set of formal series solutions $s_i(x)$ of the form

$$(1) \qquad s_i(x) = e^{Q_i(x)} x^{r_i} \sigma_i(x), \qquad Q_i(x) = \sum_{\nu=0}^{l_i-1} q_\nu^i x^{(l_i-\nu)/k_i} \qquad (i = 1, \cdots, n),$$

where the $\sigma_i(x)$ are of the form

$$(1a) \qquad \sigma_i(x) = \sum_{h=0}^{m_i} \log^h x \; {}_h\eta^{m_i}(x),$$

* Some of Nörlund's work is as follows. Acta Mathematica, vol. 37 (1914), pp. 327–387; *Leçons sur les Séries d'Interpolation*, Paris, 1926 (this book contains an extensive bibliography—pp. 228, 233); *Leçons sur les Equations Linéaires aux Différences Finies*, Paris, 1929.

† In this connection we shall mention *Integration linearer Differentialgleichungen durch Laplacesche Integrale und Fakultätenreihen*, Jahresbericht der Deutschen Mathematiker-Vereinigung, vol. 24 (1915), pp. 309–329; *Laplacesche Integrale, Binomialkoefficientenreihen und Gammaquotientenreihen in der Theorie der linearen Differentialgleichungen*, Mathematische Zeitschrift, vol. 21 (1924), pp. 85–95.

(1b) $$_h\eta^{m_i}(x) = \sum_{s=0}^{\infty} {}_h\eta_s^{m_i} x^{-s/k_i} \qquad (h = 0, 1, \cdots, m_i);$$

here l_i, m_i, k_i are integers ($m_i \geqq 0$; $k_i = r_i' p$; integer $r_i' \geqq 1$).

The formal series can be all arranged in *logarithmic groups*. The exponential factors,

$$e^{Q(x)} x^r,$$

of the series belonging to a particular group are the same.* The elements of such a group can be so ordered,

(2) $$e^{Q(x)} x^r \sigma_{i_1+1}(x),\ e^{Q(x)} x^r \sigma_{i_1+2}(x),\ \cdots,\ e^{Q(x)} x^r \sigma_{i_1+j}(x),$$

that $m_{i_1+1} = 0$, $m_{i_1+2} = 1$, $\cdots$, $m_{i_1+j} = j-1$. *A series of the type of* (1a) *will be termed a σ-series.*

Horn assumes that all the roots of the characteristic equation, corresponding to (A), are simple. Under this supposition all the formal series will be normal; moreover, no σ-series factor will contain any logarithms (that is, for each formal solution $m_i = 0$). It will be convenient to state his result in a form slightly different from his. *By means of a transformation of the type $x^k = z$, where k is suitably chosen, the equation* (A) *is brought to the form of a new equation* (A_1), the latter equation differing from the original one in the value of the integer p, while $d_{0,0} \neq 0$ and the coefficients of the new equation contain no positive (integral or fractional) powers of z. For simplicity the notation originally introduced for the equation (A) will be maintained for the modified equation (A_1). The connection between the results valid for (A) and those valid for (A_1) is obvious. Now, under Horn's hypothesis equation (A_1) possesses a set of n (linearly independent) solutions of the form

(3) $$e^{Q_i(x)} x^{r_i} y_i(x) \qquad (i = 1, 2, \cdots, n).$$

Here the $Q_i(x)$ are given by (1) (with $k_i = p$) and the $y_i(x)$ are of the form

(4) $$y(x) = {}_0\eta_0^0 + \sum_{w=1}^{p} x^{1-w/p}\, {}_0\eta_w^0(x),$$

where the ${}_0\eta_w^0(x)$ are convergent factorial series of the type

(4a) $$_0\eta_w^0(x) = \sum_{s=0}^{\infty} \frac{{}_0a_{w:s}^0}{x(x-\gamma)\cdots(x-s\gamma)},$$

* However, the values of r, associated with the same group, may differ by rational fractions.

$|\gamma|$ is sufficiently great, is the same for all solutions and $\angle\gamma$ is allowed to have any value except certain ones, depending on the roots of the characteristic equation.* Moreover, the series (4a) all converge in a certain half plane. Formally, of course, solutions (3) are compatible with a set of (possibly divergent) formal series solutions.

If the characteristic equation of (A_1) has a simple root, then to this root there will correspond a convergent solution of the form ((3), (4), (4a)). To every simple root of the characteristic equation of (A_1) there corresponds such a convergent solution. This fact, although not demonstrated by Horn, is a rather easy consequence of his work. However, we shall proceed to prove a result reaching much further. Before formulating our objective more precisely the distinction will be first drawn between "normal" and "anormal" formal series solutions. *A formal solution* $s_i(x)$ *of* (A_1), *as given by* (1), (1a), (1b), *is said to be normal when* $k_i = p$ (cf. (1b); the coefficients of (A_1) are in powers of $x^{1/p}$). *If this is not the case* (that is, when the integer r_i', of (1b), is greater than unity) *a formal series* $s_i(x)$ *will be said to be anormal.* A formal solution of (A) will be normal or anormal according to the nature of the corresponding solutions of (A_1).

In the sequel, unless stated to the contrary, the equation (A) will be taken in the form (A_1). Consider a root ρ of multiplicity $\phi(\geqq 1)$. It may happen that the ϕ formal series solutions, corresponding to this root, are all anormal; in some cases some of the ϕ formal solutions, associated with such a multiple root, are normal while others are anormal. *The third alternative—the one from now on assumed—is that all the formal solutions* (1), *belonging to the particular root under consideration, are normal. Moreover, it will be assumed that there is just one corresponding logarithmic group.* It will be shown, under this assumption, that all the formal solutions under consideration are expressible with the aid of convergent Laplace integrals, leading to convergent factorial series developments. An example in §7 will make it evident that in a more general case a result of this type need not hold.

3. Conditions for existence of formal solutions of the type specified in §2. The ϕ normal formal solutions, corresponding to a root of multiplicity ϕ, form a logarithmic group (cf. §2) and they may be written as follows:

$$s_j(x) = e^{Q(x)} x^r \sum_{h=0}^{j-1} \log^h x \; {}_h\eta^{j-1}(x) \qquad (j = 1, 2, \cdots, \phi), \tag{1}$$

where

$$ {}_h\eta^{j-1}(x) = \sum_{s=0}^{\infty} {}_h\eta_s^{j-1} x^{-s/p} \qquad (h = 0, 1, \cdots, j-1). \tag{1a}$$

* For the exact situation concerning $\angle\gamma$ cf. Horn's work on differential equations, loc. cit.

The equation (A_1) (cf. beginning of §2) will maintain its form* after the transformation

$$y(x) = e^{Q(x)} x^r \bar{y}(x).$$

Accordingly, without any loss of generality, it may be assumed that in (1)

$$e^{Q(x)} x^r \equiv 1. \tag{1b}$$

Moreover, it will be assumed, as it may be without entailing any loss of generality, that not all the

$$ {}_{j-1}\eta_0^{j-1} \qquad (j = 1, 2, \cdots, \phi) \tag{1c}$$

are zero. However, (1b) implies certain conditions on the coefficients of (A_1). In view of our purpose it will be essential to determine these conditions.

We note first that, when $h(\geqq 0)$ is an integer,

$$\frac{d^\nu}{dx^\nu} \log^h x = \sum_{\beta=0}^{h} \log^{h-\beta} x C_\beta^h [(-1)^{\nu+\beta} \nu! g_{\nu,\beta} x^{-\nu}], \tag{2}$$

where, for $\nu \geqq \beta \geqq 1$,

$$g_{\nu,\beta} = \sum_{n_1, n_2, \cdots, n_\beta} \frac{1}{n_1 n_2 \cdots n_\beta}, \tag{2a}$$

while $g_{\nu,\beta} = 0$ for $\nu < \beta$. The summation in (2a) is extended over all the positive integral values of $n_1, n_2, \cdots, n_\beta$ such that $n_1 + \cdots + n_\beta = \nu$;† moreover, $g_{0,0} = 1$ while the $g_{\nu,0}$ and the $g_{0,\nu}$ are all zero for $\nu > 0$. In view of (1b), on making use of (2), we get

$$\begin{aligned} s_j^{(k)}(x) &= \sum_{h=0}^{j-1} \sum_{\nu=0}^{k} C_\nu^k \left(\frac{d^\nu}{dx^\nu} \log^h x \right) {}_h\eta^{j-1(k-\nu)}(x) \\ &= \sum_{h=0}^{j-1} \sum_{\beta=0}^{h} \sum_{s=0}^{\infty} \sum_{\nu=0}^{k} C_\nu^k C_\beta^h C_{k-\nu}^{-s/p} (-1)^{\nu+\beta} \nu! (k-\nu)! g_{\nu,\beta} \, {}_h\eta_s^{j-1} \log^{h-\beta} x \; x^{-s/p-k}. \end{aligned}$$

Since

$$\sum_{h=0}^{j-1} \sum_{\beta=0}^{h} a_{h,\beta} l_{h-\beta} = \sum_{h=0}^{j-1} \sum_{\beta=0}^{j-1-h} a_{h+\beta,\beta} l_h, \tag{3}$$

it follows that

$$s_j^{(k)}(x) = \sum_{h=0}^{j-1} \sum_{s=0}^{\infty} s_{h,s}^{k,j} \log^h x \; x^{-s/p-k}, \tag{4}$$

---

* In particular, the coefficients of the equation will still be in powers of $x^{1/p}$.

† Formula (2) can be proved, for instance, on the basis of the multinomial theorem of algebra. The C_β^h denote binomial coefficients.

where

$$s_{h,s}^{k,j} = \sum_{\beta=0}^{j-1-h} \sum_{\nu=0}^{k} C_\nu^k C_\beta^{h+\beta} C_{k-\nu}^{-s/p} (-1)^{\nu+\beta} \nu!(k-\nu)! g_{\nu,\beta\ h+\beta}\eta_s^{j-1}. \tag{4a}$$

As a consequence of the easily verifiable relation

$$\sum_{k=0}^{n} \sum_{\sigma=0}^{\infty} b_{k,\sigma} x^{-\sigma/p-k} = \sum_{\lambda=0}^{\infty} \sum_{w=0}^{p-1} x^{-(\lambda p+w)/p} \sum_{k=0(\leqq n)}^{\lambda} b_{k,(\lambda-k)p+w} \tag{5}$$

it follows from (4) that formally

$$\begin{aligned} L(s_j(x)) &\equiv \sum_{k=0}^{n} \sum_{\lambda=0}^{\infty} d_{n-k,\lambda} x^{-\lambda/p} \sum_{h=0}^{j-1} \sum_{s=0}^{\infty} s_{h,s}^{k,j} \log^h x \ x^{-(s/p+k)} \\ &\equiv \sum_{h=0}^{j-1} \log^h x \sum_{\lambda=0}^{\infty} \sum_{w=0}^{p-1} W_{h:\lambda p+w}^{j} x^{-(\lambda p+w)/p} \qquad (j = 1, 2, \cdots, \phi). \end{aligned} \tag{6}$$

Here

$$\begin{aligned} W_{h:\lambda p+w}^{j} &= \sum_{k=0(\leqq n)}^{\lambda} \sum_{s=0}^{(\lambda-k)p+w} s_{h,s}^{k,j} d_{n-k,(\lambda-k)p+w-s} \\ &= \sum_{k=0(\leqq n)}^{\lambda} \sum_{s=0}^{(\lambda-k)p+w} \sum_{\beta=0}^{j-1-h} \sum_{\nu=0}^{k} C_\nu^k C_\beta^{h+\beta} C_{k-\nu}^{-s/p} (-1)^{\nu+\beta} \nu!(k-\nu)! g_{\nu,\beta} \\ &\qquad \cdot d_{n-k,(\lambda-k)p+w-s\ h+\beta}\eta_s^{j-1}. \end{aligned} \tag{7}$$

In view of the assumed existence of ϕ formal solutions (1) it is inferred that the equations

$$W_{h:\lambda p+w}^{j} = 0 \tag{7a}$$

$[\lambda=0, 1, \cdots; h=j-1, j-2, \cdots, 0; w=0, 1, \cdots, p-1; j=1, 2, \cdots, \phi]$

are necessarily formally solvable for the ${}_h\eta_s{}^{j-1}$. Now

$$\sum_{k=0}^{\lambda} \sum_{s=0}^{(\lambda-k)p+w} a_{k,s}\eta_s = \sum_{s=0}^{\lambda p+w} \sum_{k=0}^{k(s)} a_{k,s}\eta_s,$$

where $k(s)=\lambda$ $(s=0, 1, \cdots, w)$, $k(s)=\lambda-1$ $(s=w+1, \cdots, p+w)$, $k(s)=\lambda-2$ $(s=p+w+1, \cdots, 2p+w)$, $\cdots$, $k(s)=0$ $(s=(\lambda-1)p+w+1, \cdots, \lambda p+w)$. This relation leads to the following:

$$\begin{aligned} \sum_{k=0}^{\lambda} \sum_{s=0}^{(\lambda-k)p+w} a_{k,s}\eta_s &= \sum_{s=0}^{w} \sum_{k=0}^{\lambda} a_{k,s}\eta_s \\ &\quad + \sum_{m=1}^{\lambda} \sum_{s=s'}^{mp+w} \sum_{k=0}^{\lambda-m} a_{k,s}\eta_s \qquad [s' = (m-1)p + w + 1]. \end{aligned} \tag{8}$$

In (7) replace h by $j-H$ $(H=1, 2, \cdots, j)$. Application of (8) to (7) will then yield

$$
\begin{aligned}
W^{j}_{j-H:\lambda p+w} \equiv J^{H-1,0}_{\lambda:w}\; {}_{j-1}\eta_0^{j-1} &+ \sum_{s=0}^{w}\sum_{\beta=0}^{H-1} C_\beta^{j-H+\beta}\; J^{\beta,s}_{\lambda:w-s}\; {}_{j-H+\beta}\eta_s^{j-1} \\
&+ \sum_{m=1}^{\lambda}\sum_{s=s'}^{mp+w}\sum_{\beta=0}^{H-1} C_\beta^{j-H+\beta}\; J^{\beta,s}_{\lambda-m:mp+w-s}\; {}_{j-H+\beta}\eta_s^{j-1} = 0
\end{aligned} \tag{9}
$$

$$[\lambda = 0, 1, \cdots; H = 1, 2, \cdots, j; w = 0, 1, \cdots, p-1; s' \text{ as in (8)}].$$

Here and in the sequel

$$\sum_{s=0}^{w}\sum_{\beta=0}^{H-1} b_{s,\beta} = \sum_{s=0}^{w}\sum_{\beta=0}^{H-1} b_{s,\beta} - b_{0,H-1}; \tag{9a}$$

moreover,

$$J^{\beta,s}_{\sigma:w} = \sum_{k=\beta}^{\sigma}\sum_{\nu=\beta}^{k} C_\nu^{k} C_{k-\nu}^{-s/p}(-1)^{\nu+\beta}\nu!(k-\nu)!g_{\nu,\beta}\, d_{n-k,(\sigma-k)p+w} \tag{9b}$$

$$(0 \leqq \beta \leqq \sigma;\ s \geqq 0;\ 0 \leqq w \leqq p-1).$$

It is observed that (9b) defines all the $J^{\beta,s}_{\sigma:w}$ occurring in the second members of (9).

From equations $(9; \lambda=0; H=1; w=0, 1, \cdots, p-1)$ we find that

$$J^{0,0}_{0:w} = 0 \qquad (w = 0, 1, \cdots, p-1); \tag{10}$$

this, however, as follows from (9b), is equivalent to

$$d_{n,w} = 0 \qquad (w = 0, 1, \cdots, p-1). \tag{10a}$$

At this point it will be convenient to introduce the

Definition. *A number $d_{i,j}$ will be said to be of index σ if $i=n-k$ and $j=(\sigma-k)p+w$, where $0 \leqq k \leqq \sigma$ and $0 \leqq w \leqq p-1$.*

Thus the implication of (10) is that all the $d_{i,j}$ of index zero are zero. We note that, as a consequence of the original hypothesis concerning existence of solutions, the characteristic equation of the differential equation (A_1),

$$E(\rho) \equiv \sum_{k=0}^{n} d_{n-k,0}\,\rho^k = 0, \tag{11}$$

has a root $\rho=0$ whose multiplicity is precisely ϕ. Thus

$$d_{n,0} = d_{n-1,0} = \cdots = d_{n-\phi+1,0} = 0,\quad d_{n-\phi,0} \neq 0. \tag{11a}$$

In particular, then, it is to be noted that *not all the $d_{i,j}$ of index ϕ are zero.* Suppose

$$(12)\qquad J^{0,0}_{\sigma:w} = J^{1,0}_{\sigma:w} = \cdots = J^{\sigma,0}_{\sigma:w} = 0 \qquad (w = 0, 1, \cdots, p-1)$$

for $\sigma=0, 1, \cdots, \lambda-1$ $(1\leqq\lambda\leqq\phi-1)$. From (10) it is seen that (12) is true for $\lambda=1$. We have, by (9b),

$$(13)\qquad J^{\beta,0}_{\lambda:v} = \sum_{k=\beta}^{\lambda} (-1)^{k+\beta} k! g_{k,\beta}\, d_{n-k,(\lambda-k)p+v}.$$

Accordingly, by (12), in view of the relations $J^{\sigma,0}_{\sigma:w}=0$, it follows that

$$(14)\qquad d_{n-\sigma,w} = 0 \qquad (\sigma = 0, 1, \cdots, \lambda-1;\ w = 0, 1, \cdots, p-1);$$

furthermore the relations

$$J^{\sigma-1,0}_{\sigma:w} = 0$$

yield, by virtue of (14),

$$d_{n-(\sigma-1),p+w} = 0 \qquad (\sigma = 1, 2, \cdots, \lambda-1;\ w = 0, 1, \cdots, p-1).$$

From

$$J^{\sigma-2,0}_{\sigma:w} = 0 \qquad (\sigma = 2, \cdots, \lambda-1;\ w = 0, \cdots, p-1),$$

by (14) and (14a),

$$d_{n-(\sigma-2),2p+w} = 0 \qquad (\sigma = 2, \cdots, \lambda-1;\ w = 0, \cdots, p-1).$$

On using the relations

$$J^{\sigma-H,0}_{\sigma:w} = 0 \qquad (\sigma = H, H+1, \cdots, \lambda-1;\ w = 0, 1, \cdots, p-1)$$

in succession for $H=0, 1, \cdots, \lambda-1$ it follows by induction that

$$(14a)\qquad d_{n-(\sigma-H),Hp+w} = 0$$
$$(\sigma = H, H+1, \cdots, \lambda-1;\ w = 0, 1, \cdots, p-1)$$

for $H=0, 1, \cdots, \lambda-1$. The subscripts in (14a) can also be considered as extending over the values

$$H = 0, \cdots, \sigma; \qquad \sigma = 0, 1, \cdots, \lambda-1; \qquad w = 0, 1, \cdots, p-1.$$

Consequently, on letting in (14a) $H=\sigma-k$, (14a) is seen to be equivalent to

$$(15)\qquad d_{n-k,(\sigma-k)p+w} = 0 \qquad (k = 0, 1, \cdots, \sigma;\ w = 0, 1, \cdots, p-1),$$

where $\sigma=0, 1, \cdots, \lambda-1$. *That is,* (12) *implies that all the* $d_{i,j}$ *of indices* $0, 1, \cdots, \lambda-1$ *are zero. The converse is also true.*

All the $d_{i,j}$ in the second members of (9b) are of index σ; hence *a further consequence of* (12) *would be*

$$J^{\beta,s}_{\sigma:w} = 0 \tag{16}$$

$$(\beta=0, 1, \cdots, \sigma;\ w=0, 1, \cdots, p-1;\ \sigma=0, 1, \cdots, \lambda-1;\ s=0, 1, \cdots).$$

In view of (16) from the equations

$$W^{j}_{j-1:\lambda p} = W^{j}_{j-2:\lambda p} = \cdots = W^{j}_{j-\lambda-1:\lambda p} = 0 \tag{17}$$

we obtain in succession

$$J^{0,0}_{\lambda:0} = J^{1,0}_{\lambda:0} = \cdots = J^{\lambda,0}_{\lambda:0} = 0. \tag{17a}$$

Suppose now, more generally, that

$$J^{\beta,0}_{\lambda:v} = 0 \qquad (\beta = 0, 1, \cdots, \lambda) \tag{17b}$$

for $v=0, 1, \cdots, w-1$ $(1 \leqq w \leqq p-1)$; in (17a) the relations (17b) have been established for $w=1$. Consider the equations

$$W^{j}_{j-1:\lambda p+w} = W^{j}_{j-2:\lambda p+w} = \cdots = W^{j}_{j-(\lambda+1):\lambda p} = 0. \tag{18}$$

By (16) the numbers $J^{\beta,s}_{\lambda-m:mp+w-s}$, occurring in the second members of (9), are all zero; moreover, in consequence of (17b) the $J^{\beta,s}_{\lambda:w-s}$ (in (9)) are all zero for $w-s \leqq w-1$. Thus, equations (18) are of the form*

$$W^{j}_{j-H:\lambda p+w} \equiv J^{H-1,0}_{\lambda:w}\ {}_{j-1}\eta_0^{j-1} + \sum_{\beta=0}^{H-2} C_\beta^{j-H+\beta}\ J^{\beta,0}_{\lambda:w}\ {}_{j-H+\beta}\eta_0^{j-1} = 0. \tag{18a}$$

On using (18a) in succession for $H=1, 2, \cdots, \lambda+1$ it is found that

$$J^{H-1,0}_{\lambda:w} = 0 \qquad (H = 1, 2, \cdots, \lambda+1);$$

that is, relations (17b) necessarily hold for $v=w$, if they hold for $v=0, 1, \cdots, w-1$ $(1 \leqq w \leqq p-1)$. Thus, by induction, (12) has been established for $\sigma=0, 1, \cdots, \lambda$. This fact completes an induction in a larger sense; that is, (12) is seen to be true for $\lambda=1, 2, \cdots, \phi-1$. Just as the two italicized statements in connection with (15) and (16) had been established on the basis of (12) (as originally formulated), we now conclude that *all the* $d_{i,j}$ *of indices* $0, 1, \cdots, \phi-1$ *are zero and that*

---

* In (18a) and throughout the paper $\sum_{i=\alpha}^{\beta}=0$ whenever $\beta<\alpha$.

$$J_{\sigma:w}^{\beta,s} = 0 \tag{19}$$

$$(\beta = 0, 1, \cdots, \sigma;\ w = 0, \cdots, p-1;\ \sigma = 0, \cdots, \phi - 1;\ s = 0, 1, \cdots).$$

It is noted that, in consequence of (19), (9) holds for ($\lambda = 0, 1, \cdots, \phi-1$; $H = 1, \cdots, j$; $w = 0, 1, \cdots, p-1$; $j = 1, \cdots, \phi$), while for these values of the subscripts and superscripts the equalities (9) do not yield any information concerning the coefficients of the formal solutions, whose existence has been postulated. It remains to consider the equations (9) for $\lambda \geqq \phi$. On account of (19), in (9), $m \leqq \lambda - \phi$. For $\lambda = \phi$ these equations give

$$J_{\phi:0}^{0,w}\ {}_{j-H}\eta_{w}^{j-1} = -\sum_{s=0}^{w}\sum_{\beta=0}^{H-1} C_{\beta}^{j-H+\beta}\ J_{\phi:w-s}^{\beta,s}\ {}_{j-H+\beta}\eta_{s}^{j-1} \tag{20}$$

$$(H = 1, 2, \cdots, j;\ w = 0, 1, \cdots;\ (s, \beta) \neq (w, 0)).$$

In particular,

$$J_{\phi:0}^{0,0}\ {}_{j-1}\eta_{0}^{j-1} = 0 \qquad (j = 1, \cdots, \phi),$$

so that necessarily

$$J_{\phi:0}^{0,0} = d_{n,\phi p} = 0. \tag{21}$$

For $\lambda > \phi$, (9) yields the relations

$$\begin{aligned} J_{\phi:0}^{0,(\lambda-\phi)p+w}\ {}_{j-H}\eta_{(\lambda-\phi)p+w}^{j-1} = &-\sum_{s=0}^{w}\sum_{\beta=0}^{H-1} C_{\beta}^{j-H+\beta}\ J_{\lambda:w-s}^{\beta,s}\ {}_{j-H+\beta}\eta_{s}^{j-1} \\ &-\sum_{m=1}^{\lambda-\phi}\sum_{s=s'}^{mp+w}\sum_{\beta=0}^{H-1} C_{\beta}^{j-H+\beta}\ J_{\lambda-m:mp+w-s}^{\beta,s}\ {}_{j-H+\beta}\eta_{s}^{j-1} \end{aligned} \tag{22}$$

$$\begin{aligned}[\lambda = \phi + 1, \phi + 2, \cdots;\ H = 1, 2, \cdots, j;\ w = 0, 1, \cdots, p - 1;& \\ j = 1, 2, \cdots, \phi;\ (s, \beta) \neq ((\lambda - \phi)p + w, 0);\ s' \text{ as in } (8)].&\end{aligned}$$

The equations (20), (22) are solvable in the following order:

$$\lambda = \phi \begin{cases} H = 1;\ w = 0, 1, \cdots, p-1 \\ H = 2;\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ H = j;\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \end{cases} \Bigg|\ ;$$

$$\lambda = \phi + 1 \begin{cases} H = 1;\ w = 0, 1, \cdots, p-1 \\ H = 2;\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ H = j;\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \end{cases} \Bigg|\ ; \quad \lambda = \phi + 2, \cdots. \tag{23}$$

In (20) and (22), by (21) and (9b),

$$J_{\phi:0}^{0,(\lambda-\phi)p+w} = \sum_{k=1}^{\phi} C_k^{-((\lambda-\phi)p+w)/p} k! d_{n-k,(\phi-k)p} \tag{24}$$

$$(\lambda = \phi, \phi + 1, \cdots ; w = 0, 1, \cdots , p - 1).$$

Since $d_{n-\phi,0} \neq 0$ *only a finite number of the left members in* (24) *may vanish.* Whenever, for some w $(0 \leqq w \leqq p-1)$ and for some λ $(\lambda \geqq 0)$, a number $J_{\phi:0}^{0,(\lambda-\phi)p+w} = 0$ the corresponding ${}_{j-H}\eta_{(\lambda-\phi)p+w}^{j-1}$ is left undefined. On the other hand, for such a pair of values (w, λ), the second member in (22) or, if $\lambda = \phi$, the second member in (20) will be necessarily zero. The ${}_{j-H+\beta}\eta_s^{j-1}$ involved in such a member have known values or some of them may have been previously left undefined. The totality of relationships of such an origin, finite in number, implies certain conditions on the η^{j-1} and on the coefficients of (A_1); the latter conditions are necessarily satisfied in view of the assumed existence of solutions of stated type. Some of the η^{j-1} may be arbitrary.* The precise nature of the conditions implied by the vanishing of (24) is immaterial for our purposes.

LEMMA 1. *Consider a root of multiplicity* ϕ, *of the characteristic equation associated with the differential equation* (A_1). *In order that, corresponding to this root, there should exist a linearly independent set of* ϕ *formal solutions of type* (1), (1a), (1b) *the following conditions are necessary and sufficient.*

(i) $$d_{n-\phi,0} \neq 0, \qquad d_{n,\phi p} = 0.$$

(ii) *All the* $d_{i,j}$, *whose indices* (*cf. Definition*) *are* $0, 1, \cdots, \phi-1$, *are zero.*

(iii) *If any of the* $J_{\phi:0}^{0,(\lambda-\phi)p+w}$, *defined by* (24), *are zero* (*there may be only a finite number of such* J) *then the* $d_{i,j}$ *satisfy conditions implied by the vanishing of the corresponding second members of* (22) *or* (20).

4. **The mixed system of differential equations.** In the sequel, unless stated to the contrary, it will be assumed that by means of the transformation of the type specified in the beginning of §3 the differential equation (A_1) has been brought to such a form that there exist ϕ (linearly independent) formal solutions (1; §3), for which (1b; §3) holds. These solutions may be written as follows:

$$s_j(x) = \sum_{h=0}^{j-1} \log^h x \left({}_h\eta_0^{j-1} + \sum_{w=1}^{p} x^{(p-w)/p} {}_h\eta_w^{j-1}(x) \right) \qquad (j = 1, \cdots, \phi), \tag{1}$$

---

* In general, whenever some of the numbers (24) are zero, associated with the root, under consideration, of the characteristic equation there will exist more than one formal series solution of (A_1), not involving logarithms.

where

$$\text{(1a)}\qquad {}_h\eta_w^{j-1}(x) = \sum_{\lambda=0}^{\infty} {}_h\eta^{j-1}_{\lambda p+w}\, x^{-\lambda-1} \qquad (h = 0, 1, \cdots, j-1;\ w = 1, \cdots, p).$$

On the other hand, the coefficients $d_{n-k}(x)$ of (A_1) may be expressed in the form

$$\text{(2)}\qquad d_{n-k}(x) = d_{n-k,0} + \sum_{\nu=1}^{p} x^{(p-\nu)/p}\, d_{n-k,\nu}(x),$$

$$\text{(2b)}\qquad d_{n-k,\nu}(x) = \sum_{\lambda=0}^{\infty} d_{n-k,\lambda p+\nu}\, x^{-\lambda-1} \qquad (k = 0, 1, \cdots, n).$$

For the purposes at hand it will be essential to establish a "mixed" linear differential system, whose coefficients are in negative integral powers of x and which are formally satisfied by the series (1a). We have

$$\text{(3)}\qquad \begin{aligned} &\frac{d^{k-m}}{dx^{k-m}}\left({}_h\eta_0^{j-1} + \sum_{w=1}^{p} x^{(p-w)/p}\, {}_h\eta_w^{j-1}(x)\right) \\ &\quad = {}_h\eta_0^{j-1(k-m)} + \sum_{w=1}^{p}\sum_{\delta=0}^{k-m} C_\delta^{k-m} C_\delta^{(p-w)/p}\,\delta!\, {}_h\eta_w^{j-1(k-m-\delta)}(x)\, x^{(p-w)/p-\delta}. \end{aligned}$$

By (3) and by (2; §3) it follows from (1) that

$$\text{(4)}\qquad \begin{aligned} s_j^{(k)}(x) = {} & \sum_{h=0}^{j-1}\sum_{m=0}^{k} C_m^k \sum_{\beta=0}^{h} \log^{h-\beta} x\; C_\beta^h (-1)^{m+\beta} m!\, g_{m,\beta}\, x^{-m} \\ & \times \left[{}_h\eta_0^{j-1(k-m)} + \sum_{w=1}^{p}\sum_{\delta=0}^{k-m} C_\delta^{k-m} C_\delta^{(p-w)/p}\,\delta!\, {}_h\eta_w^{j-1(k-m-\delta)}(x)\, x^{(p-w)/p-\delta}\right]. \end{aligned}$$

This may be written in the form

$$\text{(4a)}\qquad s_j^{(k)}(x) = \sum_{h=0}^{j-1} \log^h x\,[S_1^{k,j} + S_2^{k,j}],$$

where

$$\text{(4b)}\qquad S_1^{k,j} = \sum_{\beta=0}^{j-h-1} C_\beta^{h+\beta} (-1)^{k+\beta} k!\, g_{k,\beta}\, x^{-k},$$

$$\text{(4c)}\qquad \begin{aligned} S_2^{k,j} = {} & \sum_{\beta=0}^{j-h-1}\sum_{w=1}^{p}\sum_{m=0}^{k}\sum_{\delta=0}^{m} C_{m-\delta}^k C_\beta^{h+\beta} C_\delta^{k-m+\delta} C_\delta^{(p-w)/p} \\ & \times (m-\delta)!\,\delta!\,(-1)^{m-\delta+\beta} g_{m-\delta,\beta}\; {}_{h+\beta}\eta_w^{j-1(k-m)}(x)\, x^{-m} x^{(p-w)/p}. \end{aligned}$$

Thus

$$
\begin{aligned}
(5)\qquad L(s_j) \equiv \sum_{h=0}^{j-1} \log^h x \sum_{k=0}^{n} \Big[& S_1^{k,j} d_{n-k,0} + S_1^{k,j} \sum_{\nu=1}^{p} x^{(p-\nu)/p} d_{n-k,\nu}(x) \\
& + S_2^{k,j} d_{n-k,0} + S_2^{k,j} \sum_{\nu=1}^{p} x^{(p-\nu)/p} d_{n-k,\nu}(x) \Big] = 0.
\end{aligned}
$$

In (5)

$$
\begin{aligned}
(6)\qquad & S_1^{k,j} \sum_{\nu=1}^{p} x^{(p-\nu)/p} d_{n-k,\nu}(x) \\
& = \sum_{w=1}^{p} \left(\sum_{\beta=0}^{j-h-1} C_\beta^{h+\beta} (-1)^{k+\beta} k!\, g_{k,\beta}\, d_{n-k,w}(x) x^{-k} \right) x^{(p-w)/p},
\end{aligned}
$$

and, by virtue of the relation

$$
(7)\qquad \sum_{w=1}^{p} \sum_{\nu=1}^{p} a_{w,\nu}\, x^{(2p-w-\nu)/p} = \sum_{w=1}^{p} x^{(p-w)/p} \left[\sum_{\nu=1}^{w-1} x a_{w-\nu,\nu} + \sum_{\nu=w}^{p} a_{p+w-\nu,\nu} \right],
$$

we have

$$
\begin{aligned}
& S_2^{k,j} \sum_{\nu=1}^{p} x^{(p-\nu)/p} d_{n-k,\nu}(x) \\
& = \sum_{w=1}^{p} x^{(p-w)/p} \Big[\sum_{\nu=1}^{w-1} \sum_{m=0}^{k} \sum_{\beta=0}^{j-h-1} \sum_{\delta=0}^{m} C_{m-\delta}^{k} C_\beta^{h+\beta} C_\delta^{k-m+\delta} C_\delta^{(p-w+\nu)/p} (m-\delta)! \\
(8)\qquad & \quad \times \delta! (-1)^{m-\delta+\beta} g_{m-\delta,\beta}\, d_{n-k,\nu}(x)\, {}_{h+\beta}\eta_{w-\nu}^{j-1(k-m)}\, x^{-(m-1)} \\
& \quad + \sum_{\nu=w}^{p} \sum_{m=0}^{k} \sum_{\beta=0}^{j-h-1} \sum_{\delta=0}^{m} C_{m-\delta}^{k} C_\beta^{h+\beta} C_\delta^{k-m+\delta} C_\delta^{\nu-w} (m-\delta)!\, \delta! (-1)^{m-\delta+\beta} \\
& \quad \times g_{m-\delta,\beta}\, d_{n-k,\nu}(x)\, {}_{h+\beta}\eta_{w+p-\nu}^{j-1(k-m)}(x) x^{-m} \Big].
\end{aligned}
$$

By (4b), (4c), (6) and (8) we obtain from (5)

$$
(9)\qquad L(s_j(x)) \equiv \sum_{h=0}^{j-1} \sum_{w=1}^{p} \log^h x \;\; x^{(p-w)/p} W_{h,w}^{j-1} = 0.
$$

Here

$$
(9a)\qquad W_{h,w}^{j-1} = W_1^{h,w,j-1} + W_2^{h,w,j-1} + W_3^{h,w,j-1} + W_4^{h,w,j-1}
$$

where

$$(9b)\quad \begin{aligned} W_1^{h,w,j-1} &= \xi(w)\sum_{k=0}^{n}\sum_{\beta=0}^{j-h-1} C_\beta^{h+\beta}(-1)^{k+\beta}k!\,g_{k,\beta}d_{n-k,0}x^{-k} \\ &+ \sum_{k=0}^{n}\sum_{\beta=0}^{j-h-1} C_\beta^{h+\beta}(-1)^{k+\beta}k!\,g_{k,\beta}\,d_{n-k,w}(x)x^{-k} \\ &\qquad (\xi(w) = 0 \text{ for } w \neq p;\ \xi(p) = 1), \end{aligned}$$

$$(9c)\quad \begin{aligned} W_2^{h,w,j-1} &\equiv \sum_{k=0}^{n}\sum_{m=0}^{n-k}\sum_{\beta=0}^{j-h-1}\sum_{\delta=0}^{m} C_{m-\delta}^{k+m}C_\beta^{h+\beta}C_\delta^{k+\delta}C_\delta^{(p-w)/p}(m-\delta)! \\ &\times \delta!(-1)^{m-\delta+\beta}g_{m-\delta,\beta}\,d_{n-k-m,0}\ {}_{h+\beta}\eta_w^{j-1(k)}(x)x^{-m}, \end{aligned}$$

$$(9d)\quad \begin{aligned} W_3^{h,w,j-1} &\equiv \sum_{k=0}^{n}\sum_{m=0}^{n-k}\sum_{\zeta=1}^{w-1}\sum_{\beta=0}^{j-h-1}\sum_{\delta=0}^{m} C_{m-\delta}^{k+m}C_\beta^{h+\beta}C_\delta^{k+\delta}C_\delta^{(p-\zeta)/p} \\ &\times (m-\delta)!\delta!(-1)^{m-\delta+\beta}g_{m-\delta,\beta}\,d_{n-k-m,w-\zeta}(x)\ {}_{h+\beta}\eta_\zeta^{j-1(k)}(x)x^{-(m-1)} \end{aligned}$$

and

$$(9e)\quad \begin{aligned} W_4^{h,w,j-1} &\equiv \sum_{k=0}^{n}\sum_{m=0}^{n-k}\sum_{\zeta=w}^{p}\sum_{\beta=0}^{j-h-1}\sum_{\delta=0}^{m} C_{m-\delta}^{k+m}C_\beta^{h+\beta}C_\delta^{k+\delta}C_\delta^{(p-\zeta)/p} \\ &\times (m-\delta)!\delta!(-1)^{m-\delta+\beta}g_{m-\delta,\beta}\,d_{n-k-m,w+p-\zeta}(x)\ {}_{h+\beta}\eta_\zeta^{j-1(k)}(x)x^{-m}\ . \end{aligned}$$

Now, in view of the formal facts involved, (9) implies that the ${}_h\eta_w^{j-1}(x)$ necessarily satisfy the set of equations

$$(10)\quad \begin{gathered} W_{h,w}^{j-1} = 0 \\ (h = 0, \cdots, j-1;\ w = 1, \cdots, p;\ j = 1, \cdots, \phi). \end{gathered}$$

On making use of (9a), (9b), (9c), (9d) and (9e) equations (10) are seen to be equivalent to the differential system

$$(A_2)\quad \begin{gathered} T_{h,w}^{j-1} \equiv \sum_{\beta=0}^{j-h-1}\sum_{k=0}^{n}\sum_{\zeta=1}^{p} {}_\beta a_{\zeta,k}^{h,w}(x)\ {}_{h+\beta}\eta_\zeta^{j-1(k)}(x) = g^{h,w,j}(x) \\ (h = 0, 1, \cdots, j-1;\ w = 1, \cdots, p), \end{gathered}$$

where the ${}_{h+\beta}\eta_\zeta^{j-1}(x)$ are to be regarded as variables; moreover, (A_2) is formally satisfied by the series (1a). The coefficients in (A_2) are given by

$$(11)\quad \begin{aligned} {}_\beta a_{\zeta,k}^{h,w}(x) &= \sum_{m=0}^{n-k}\sum_{\delta=0}^{m} C_{m-\delta}^{k+m}C_\beta^{h+\beta}C_\delta^{k+\delta}C_\delta^{(p-\zeta)/p}(m-\delta)!\delta! \\ &\times (-1)^{m-\delta+\beta}g_{m-\delta,\beta}\,d_{n-k-m,w-\zeta}(x)x^{-(m-1)} \qquad (\zeta = 1, \cdots, w-1), \end{aligned}$$

$$
\text{(11a)}\qquad \begin{aligned}
{}_\beta a^{h,w}_{\zeta,k}(x) &= \sum_{m=0}^{n-k}\sum_{\delta=0}^{m} C^{k+m}_{m-\delta}C^{h+\beta}_{\beta}C^{k+\delta}_{\delta}C^{(p-\zeta)/p}_{\delta}(m-\delta)!\delta!\\
&\quad\times(-1)^{m-\delta+\beta}g_{m-\delta,\beta}\,d_{n-k-m,w+p-\zeta}(x)x^{-m}\qquad(\zeta = w+1,\cdots,p),
\end{aligned}
$$

$$
\text{(11b)}\qquad \begin{aligned}
{}_\beta a^{h,w}_{w,k}(x) &= \sum_{m=0}^{n-k}\sum_{\delta=0}^{m} C^{k+m}_{m-\delta}C^{h+\beta}_{\beta}C^{k+\delta}_{\delta}C^{(p-w)/p}_{\delta}(m-\delta)!\delta!(-1)^{m-\delta+\beta}\\
&\quad\times g_{m-\delta,\beta}\,d_{n-k-m,0}\,x^{-m}+\sum_{m=0}^{n-k}\sum_{\delta=0}^{m} C^{k+m}_{m-\delta}C^{h+\beta}_{\beta}C^{k+\delta}_{\delta}C^{(p-w)/p}_{\delta}\\
&\quad\times(m-\delta)!\delta!(-1)^{m-\delta+\beta}g_{m-\delta,\beta}\,d_{n-k-m,p}(x)x^{-m}
\end{aligned}
$$

and

$$
\text{(11c)}\qquad g^{h,w,j}(x) = -W_1^{h,w,j-1}\qquad(\text{cf. (9b)}).
$$

Further calculation leads us to conclude that

$$
\text{(12)}\qquad {}_\beta a^{h,w}_{\zeta,k}(x) = \sum_{\lambda=0}^{\infty} {}_\beta a^{h,w}_{\zeta,k:\lambda}\,x^{-\lambda},\qquad g^{h,w,j}(x) = \sum_{\lambda=0}^{\infty} g_\lambda^{h,w,j}\,x^{-\lambda},
$$

where the series involved all converge in a neighborhood of infinity, and

$$
\text{(12a)}\qquad \begin{aligned}
{}_\beta a^{h,w}_{\zeta,k:\lambda} &= \sum_{\beta\leqq m\leqq n-k}^{m\leqq\lambda}\ \sum_{\delta=0}^{m-\beta} C^{k+m}_{m-\delta}C^{h+\beta}_{\beta}C^{k+\delta}_{\delta}C^{(p-\zeta)/p}_{\delta}(m-\delta)!\delta!(-1)^{m-\delta+\beta}\\
&\quad\times g_{m-\delta,\beta}\,d_{n-k-m,(\lambda-m)p+w-\zeta}\qquad(\zeta = 1,\cdots,w-1),
\end{aligned}
$$

$$
\text{(12b)}\qquad \begin{aligned}
{}_\beta a^{h,w}_{\zeta,k:\lambda} &= \sum_{\beta\leqq m\leqq n-k}^{m\leqq\lambda-1}\ \sum_{\delta=0}^{m-\beta} C^{k+m}_{m-\delta}C^{h+\beta}_{\beta}C^{k+\delta}_{\delta}C^{(p-\zeta)/p}_{\delta}(m-\delta)!\delta!\\
&\quad\times(-1)^{m-\delta+\beta}g_{m-\delta,\beta}\,d_{n-k-m,(\lambda-m)p+w-\zeta}\qquad(\zeta = w+1,\cdots,p),
\end{aligned}
$$

$$
\text{(12c)}\qquad \begin{aligned}
{}_\beta a^{h,w}_{w,k:\lambda} &= \sum_{\delta=0}^{\lambda-\beta} C^{k+\lambda}_{\lambda-\delta}C^{h+\beta}_{\beta}C^{k+\delta}_{\delta}C^{(p-w)/p}_{\delta}(\lambda-\delta)!\delta!(-1)^{\lambda-\delta+\beta}g_{\lambda-\delta,\beta}\times d_{n-k-\lambda,0}\\
&\quad+\sum_{\beta\leqq m\leqq n-k}^{m\leqq\lambda-1}\ \sum_{\delta=0}^{m-\beta} C^{k+m}_{m-\delta}C^{h+\beta}_{\beta}C^{k+\delta}_{\delta}C^{(p-w)/p}_{\delta}(m-\delta)!\delta!\\
&\quad\times(-1)^{m-\delta+\beta}g_{m-\delta,\beta}\,d_{n-k-m,(\lambda-m)p}.
\end{aligned}
$$

It is noted that, according to Definition of §3, the $d_{i,j}$ occurring in (12a) are of index $\lambda+k$, while those in (12b) are of index $\lambda+k-1$, and the index of the $d_{i,j}$ in (12c) is $\lambda+k$. Thus, by (ii) of Lemma 1,

$$
\text{(13)}\qquad {}_\beta a^{h,w}_{\zeta,k:\lambda} = 0\begin{cases}\lambda+k\leqq\phi-1 & (\zeta\leqq w);\\ \lambda+k\leqq\phi & (\zeta > w).\end{cases}
$$

It is observed, moreover, that in the summations of (12a) and (12b) we have $\beta \leqq \lambda$ and $\beta \leqq \lambda - 1$, respectively; also, in the first summation in (12c) $\lambda - \beta \geqq 0$, while in the second summation $\beta \leqq \lambda - 1$. Hence

$$(13\text{a}) \qquad {}_{\beta}\overset{h,w}{a}_{\zeta,k:\lambda} = 0 \begin{cases} \lambda \leqq \beta - 1 & (\zeta \leqq w); \\ \lambda \leqq \beta & (\zeta > w). \end{cases}$$

Finally, since in (12a) and (12b) $\beta \leqq n - k$, we conclude that

$$(13\text{b}) \qquad {}_{\beta}\overset{h,w}{a}_{\zeta,k:\lambda} = 0 \qquad (\beta + k > n;\ \zeta \neq w).$$

In particular,

$$(14) \qquad {}_{\beta}\overset{h,w}{a}_{w,k:\beta} = \overset{k+\beta}{C_{\beta}}\overset{h+\beta}{C_{\beta}}\beta! g_{\beta,\beta}\, d_{n-k-\beta,0},$$

so that

$$(14\text{a}) \qquad {}_{\beta}\overset{h,w}{a}_{w,k:\beta} = 0 \quad (k + \beta \leqq \phi - 1), \qquad {}_{\beta}\overset{h,w}{a}_{w,k:\beta} \neq 0 \qquad (k + \beta = \phi).$$

Also it is observed that

$$(15) \qquad {}_{0}\overset{h,w}{a}_{\zeta,k:0} = d_{n-k,w-\zeta} \qquad (\zeta = 1, \cdots, w - 1).$$

LEMMA 2. *Write the ϕ formal solutions* (1; §3), (1b; §3), *corresponding to a root of multiplicity ϕ of the characteristic equation associated with the differential equation* (A_1), *in the form* (1), (1a). *The formal series* (1a) (*with j fixed*) *will satisfy a "mixed" differential system* (A_2), *whose coefficients are convergent series given by* (12), (12a), (12b), (12c), (11c). *The coefficients in the series* (12) *will satisfy* (13), (13a), (13b), (14a).

5. **The corresponding system of integral equations.** In the sequel, unless stated otherwise, we shall write

$$t = |t|\, e^{i\bar{t}}$$

where $i = (-1)^{1/2}$; moreover, the integrals

$$\int_0^{\infty}, \qquad \int_0^{t}$$

will be supposed to be extended over the ray $(0, \infty)$, of angle $\bar{t}$, and the rectilinear segment $(0, t)$, respectively. The variable x will be so restricted that

$$(1) \qquad \lim_{t} |e^{tx} t^{\alpha}| = 0 \qquad (\text{every } \alpha > 0),$$

when $|t| \to \infty$ along the ray $(0, \infty)$ of angle $\bar{t}$. As is well known, we have then formally

$$a(x) = \sum_{s=1}^{\infty} a_s x^{-s} = \int_0^{\infty} \bar{a}(t) e^{tx} dt, \tag{2}$$

where

$$\bar{a}(t) = \sum_{\nu=0}^{\infty} \bar{a}_\nu t^{\nu-1}; \quad \bar{a}_\nu = \frac{(-1)^\nu a_\nu}{(\nu - 1)!}. \tag{2a}$$

Thus the series (1a; §4) are formally representable as follows:

$$ {}_h\eta_w^{j-1}(x) \left(= \sum_{\lambda=0}^{\infty} {}_h\eta_{\lambda p+w}^{j-1} x^{-\lambda-1} \right) = \int_0^{\infty} {}_h\bar{\eta}_w^{j-1}(t) e^{tx} dt, \tag{3}$$

where

$$ {}_h\bar{\eta}_w^{j-1}(t) = \sum_{\nu=0}^{\infty} {}_h\bar{\eta}_{w:\nu}^{j-1} t^\nu; \quad {}_h\bar{\eta}_{w:\nu}^{j-1} = \frac{(-1)^{\nu+1}}{\nu!} {}_h\eta_{\nu p+w}^{j-1}. \tag{3a}$$

With the part within the parenthesis deleted, (3) is to be considered as a transformation which will be applied to the differential system (A_2). The coefficients in (3a) had been previously defined by (20, 22, 23; §3). The latter relations, while useful in proving Lemma 1, are impracticable for the purpose of establishing convergence of the series (3a) (in the neighborhood of $t=0$). Their convergence, however, will be proved in the sequel with the aid of a system of integral equations and a dominant system of integral equations (§6). Besides, the integral system, to be established below, is to serve a certain other purpose (§7).

In carrying out the transformation (3) it is observed that, formally,

$$ {}_{h+\beta}\eta_\zeta^{j-1(k)}(x) = \int_0^{\infty} {}_{h+\beta}\bar{\eta}_\zeta^{j-1}(t) t^k e^{tx} dt \tag{4}$$

and that, for $\lambda \geqq 1$,

$$x^{-\lambda}\, {}_{h+\beta}\eta_\zeta^{j-1(k)}(x) = -\int_0^{\infty} \left[\int_0^t \frac{(\tau - t)^{\lambda-1}}{(\lambda - 1)!} \tau^k\, {}_{h+\beta}\bar{\eta}_\zeta^{j-1}(\tau) d\tau \right] e^{tx} dt \tag{4a}$$

provided

$$\left[e^{tx} \int_0^t t^k\, {}_h\bar{\eta}_w^{j-1}(t) dt^{(H)} \right]_0^{\infty} = 0 \tag{5}$$

$(k = 0, 1, \cdots, n;\ h = 0, \cdots, j-1;\ w = 1, \cdots, p;\ H = 1, 2, \cdots).$

In (5) the displayed integration is iterated H times.* Furthermore, by (4a) and (12; §4)

$$
(6)\qquad \begin{aligned}
{}_{\beta}a^{h,w}_{\zeta,k}(x)\ {}_{h+\beta}\eta_{\zeta}^{j-1(k)}(x) = \int_0^\infty e^{tx}\Big[&{}_{\beta}a^{h,w}_{\zeta,k:0}\,t^k\ {}_{h+\beta}\bar{\eta}_{\zeta}^{j-1}(t) \\
&- \int_0^t\Big(\sum_{\lambda=1}^{\infty} {}_{\beta}a^{h,w}_{\zeta,k:\lambda}\,\frac{(\tau-t)^{\lambda-1}}{(\lambda-1)!}\,\tau^k\Big)\,{}_{h+\beta}\bar{\eta}_{\zeta}^{j-1}(\tau)d\tau\Big]dt.
\end{aligned}
$$

By virtue of (2), we have for the last member of (A_2)

$$
(6a)\qquad g^{h,w,j}(x) = \int_0^\infty e^{tx}\bar{g}^{h,w,j}(t)dt,
$$

$$
(6b)\qquad \bar{g}^{h,w,j}(t) = \sum_{\nu=1}^{\infty}\bar{g}_{\nu}^{h,w,j}\,t^{\nu-1};\quad \bar{g}_{\nu}^{h,w,j} = \frac{(-1)^{\nu}}{(\nu-1)!}\,g_{\nu}^{h,w,j}.
$$

On account of convergence of the left members in (6a) it is observed that the left members in (6b) are entire functions. On making use of (6) and (6a) it is seen that the system (A_2) is formally satisfied if

$$
(7)\qquad \begin{aligned}
\sum_{\beta=0}^{j-h-1}\sum_{\zeta=1}^{p} {}_{\beta}b_{\zeta}^{h,w}(t)\ {}_{h+\beta}\bar{\eta}_{\zeta}^{j-1}(t) = \sum_{\beta=0}^{j-h-1}\sum_{\zeta=1}^{p}\int_0^t {}_{\beta}c_{\zeta}^{h,w}(t,\tau)\ {}_{h+\beta}\bar{\eta}_{\zeta}^{j-1}(\tau)d\tau + \bar{g}^{h,w,j}(t) \\
(h = 0, \cdots, j-1;\ w = 1, \cdots, p).
\end{aligned}
$$

Here

$$
(7a)\qquad {}_{\beta}b_{\zeta}^{h,w}(t) = \sum_{k=0}^{n} {}_{\beta}a^{h,w}_{\zeta,k:0}\,t^k,
$$

$$
(7b)\qquad {}_{\beta}c_{\zeta}^{h,w}(t,\tau) = \sum_{\lambda=1}^{\infty}\sum_{k=0}^{n} {}_{\beta}a^{h,w}_{\zeta,k:\lambda}\,\frac{(\tau-t)^{\lambda-1}}{(\lambda-1)!}\,\tau^k,
$$

and the series (7b) are entire in t and τ.

Since by (13a; §4)

$$
{}_{\beta}a^{h,w}_{\zeta,k:0} = 0 \qquad (\beta > 0;\ \zeta \leqq w), \qquad\qquad {}_{\beta}a^{h,w}_{\zeta,k:0} = 0 \qquad (\zeta > w),
$$

it follows from (7a) that

$$
(8)\qquad {}_{\beta}b_{\zeta}^{h,w}(t) = 0 \qquad (\beta > 0;\ \zeta \leqq w),
$$

$$
(8a)\qquad {}_{\beta}b_{\zeta}^{h,w}(t) = 0 \qquad (\zeta > w).
$$

---

* Those of the steps which at first can be considered as valid only in a formal sense will be finally justified. Thus, absolute convergence of the integrals (3) (for certain suitably defined functions), as well as the relations (5), will be later established.

Thus, the only functions (7a) which could possibly be not identically zero are the ${}_0b_\zeta^{h,w}(t)$ $(\zeta \leqq w)$; in view of (7a) and (15; §4) they are of the form

$$
{}_0b_\zeta^{h,w}(t) = \sum_{k=0}^{n} d_{n-k,w-\zeta}\, t^k = b^{w-\zeta}(t) \qquad (\zeta = 1, \cdots, w). \tag{9}
$$

In particular, since $d_{n-\phi,0} \neq 0$, $b^0(t) \not\equiv 0$. Thus the left members in (7) may be replaced by

$$
\sum_{\zeta=1}^{w} b^{w-\zeta}(t)\ {}_h\bar{\eta}_\zeta^{\,j-1}(t). \tag{10}
$$

The coefficients in the system (7) will be now investigated in a greater detail. From (7b) we have

$$
\begin{aligned}
{}_\beta c_\zeta^{h,w}(t, \tau) &= \sum_{s=0}^{\infty} \sum_{q=0}^{\infty} {}_\beta c_{\zeta:s,q}^{h,w}\, t^s \tau^q \\
&= \sum_{\lambda=1}^{\infty} \sum_{k=0}^{n} \sum_{r=0}^{\lambda-1} C_r^{\lambda-1} \frac{(-1)^{\lambda-1-r}}{(\lambda-1)!}\, {}_\beta a_{\zeta,k:\lambda}^{h,w}\, t^{\lambda-1-r} \tau^{r+k},
\end{aligned} \tag{11}
$$

so that

$$
{}_\beta c_{\zeta:s,q}^{h,w} = \sum_{r=0}^{q} C_r^{s+r} \frac{(-1)^s}{(s+r)!}\, {}_\beta a_{\zeta,q-r:s+r+1}^{h,w} \qquad (r \geqq q - n). \tag{11a}
$$

Thus, on noting (13; §4), it is observed that

$$
{}_\beta c_{\zeta:s,q}^{h,w} = 0 \qquad (s + q \leqq \phi - 2;\ \zeta \leqq w), \tag{12}
$$

$$
{}_\beta c_{\zeta:s,q}^{h,w} = 0 \qquad (s + q \leqq \phi - 1;\ \zeta > w)\,; \tag{12a}
$$

moreover, by virtue of (13a; §4), the summation in the right member of (11a) is extended so that $r \leqq q$ and

$$
\begin{aligned}
&r \geqq 0;\ r \geqq q - n;\ r \geqq \beta - 1 - s \qquad (\text{when } s + q \geqq \phi - 1 \text{ and } \zeta \leqq w),\\
&r \geqq 0;\ r \geqq q - n;\ r \geqq \beta - s \qquad (\text{when } s + q \geqq \phi \text{ and } \zeta > w).
\end{aligned}
$$

To establish properties at infinity we first note that, in view of the satisfied conditions of convergence,

$$
\left| {}_\beta a_{\zeta,k:\lambda}^{h,w} \right|, \qquad \left| g_\lambda^{h,w,j} \right| < R\rho^\lambda. \tag{13}
$$

These inequalities are valid for all possible values of the superscripts and subscripts.† By (13), (7b) and (6b) we shall have (exclusive of a small vicinity

† ρ^λ here denotes a power of ρ.

of $\tau=0$)

$$(14)\qquad \left| {}_{\beta}c_{\zeta}^{h,w}(t,\tau) \right| < R\rho \sum_{\lambda=1}^{\infty}\sum_{k=0}^{n} \frac{(\rho\,|\,t-\tau\,|\,)^{\lambda-1}}{(\lambda-1)!}\,|\,\tau\,|^{k} < R''\,|\,\tau\,|^{n} e^{\rho|t-\tau|}$$

and

$$(14\text{a})\qquad \left| \bar{g}^{h,w,j}(t) \right| < R\rho \sum_{\nu=1}^{\infty} \frac{(\rho\,|\,t\,|\,)^{\nu-1}}{(\nu-1)!} < R'' e^{\rho|t|}.$$

Consider the coefficients in the left members of (7) (cf. (10)). It is noted that, since the $d_{n-k,w-\zeta}$ ($\zeta=1,\cdots,w$) in (9) are of index k, application of Lemma 1 will yield the result

$$(15)\qquad b^{w-\zeta}(t) = t^{\phi} d^{w-\zeta}(t),\quad d^{w-\zeta}(t) = d_{n-\phi,w-\zeta} + d_{n-\phi-1,w-\zeta}t + \cdots + d_{0,w-\zeta}t^{n-\phi} \qquad (\zeta = 1,\cdots,w).$$

On the other hand, since $d_{n-\phi,0}\neq 0$,

$$(15\text{a})\qquad \frac{t^{\phi}}{b^{(0)}(t)} = \frac{1}{d_{n-\phi,0}+\cdots+d_{0,0}t^{n-\phi}} = d(t) = d_0 + d_1 t + \cdots \qquad (d_0 \neq 0,\ \neq \infty).$$

Let P denote the complex t-plane excluding small sectors, each with vertex at $t=0$ and containing the poles of $d(t)$ (15a) *in their interiors.*

From (15) and (15a) it follows that

$$(16)\qquad |\,d(t)\,| < d\,|\,t\,|^{-n+\phi},$$

$$(16\text{a})\qquad |\,d^{w-\zeta}(t)\,| < d\,|\,t\,|^{n-\phi} \qquad (\zeta = 1,\cdots,w)$$

when t is in P, exclusive of a small vicinity of $t=0$.

We shall now proceed to derive a modified integral system. From (7), in view of the notation introduced in (15) and (15a),

$$(17)\qquad \begin{aligned} t^{\phi}\,{}_{h}\bar{\eta}_{w}^{j-1}(t) = &-\sum_{\sigma=1}^{w-1} d(t)b^{w-\zeta}(t)\,{}_{h}\bar{\eta}_{\sigma}^{j-1}(t) + d(t)\bar{g}^{h,w,j}(t) \\ &+\sum_{\beta=0}^{j-h-1}\sum_{\zeta=1}^{p}\int_{0}^{t} d(t)\,{}_{\beta}c_{\zeta}^{h,w}(t,\tau)\,{}_{h+\beta}\bar{\eta}_{\zeta}^{j-1}(\tau)d\tau \end{aligned} \qquad (h = 0,\cdots,j-1;\ w = 1,\cdots,p).$$

It will be demonstrated that (17) can be brought to the form

$$(\text{A}_3)\qquad t^{\phi}\,{}_{h}\bar{\eta}_{w}^{j-1}(t) = \sum_{\beta=0}^{j-h-1}\sum_{\zeta=1}^{p}\int_{0}^{t} {}_{\beta}\overset{*}{c}{}_{\zeta}^{h,w}(t,\tau)\,{}_{h+\beta}\bar{\eta}_{\zeta}^{j-1}(\tau)d\tau + \overset{*}{g}{}^{h,w,j}(t) \qquad (h = 0,\cdots,j-1;\ w = 1,\cdots,p).$$

It is noted first that from $(17; w=1)$ the equations $(A_3; w=1)$ are obtained, with

$$(18) \qquad {}_{\beta}^{*}c_{\zeta}^{h,1}(t, \tau) = d(t)\, {}_{\beta}c_{\zeta}^{h,1}(t, \tau),$$

$$(18a) \qquad {}^{*}g^{h,1,j}(t) = d(t)\bar{g}^{h,1,j}(t).$$

Suppose that, for $\sigma = 1, 2, \cdots, w-1$ $(2 \leqq w \leqq p)$,

$$(19) \qquad t^{\phi}\, {}_{h}\bar{\eta}_{\sigma}^{j-1}(t) = \sum_{\beta=0}^{j-h-1} \sum_{\zeta=1}^{p} \int_{0}^{t} {}_{\beta}^{*}c_{\zeta}^{h,\sigma}(t, \tau)\, {}_{h+\beta}\bar{\eta}_{\zeta}^{j-1}(\tau) d\tau + {}^{*}g^{h,\sigma,j}(t).$$

Substitution of (19) in (17) will result in equations of form (A_3), where

$$(20) \qquad {}_{\beta}^{*}c_{\zeta}^{h,w}(t, \tau) = d(t)\, {}_{\beta}c_{\zeta}^{h,w}(t, \tau) - d(t) d^{w-\zeta}(t) \sum_{\sigma=1}^{w-1} {}_{\beta}^{*}c_{\zeta}^{h,\sigma}(t, \tau),$$

$$(20a) \qquad {}^{*}g^{h,w,j}(t) = d(t)\bar{g}^{h,w,j}(t) - \sum_{\sigma=1}^{w-1} d(t) d^{w-\zeta}(t)\, {}^{*}g^{h,\zeta,j}(t).$$

Relations (20), (20a) have been established for $w=1$ in (18), (18a). Thus, it follows by induction that (7) is equivalent to the system (A_3), where the coefficients are defined in terms of those of (7) by means of the recursion relations (20), (20a).

In view of (15a), of the nature of the coefficients in (7) and of (20) and (20a) it is concluded that the coefficients in (A_3) are meromorphic functions (in t), whose t-poles are at the non-zero roots of the characteristic equation of (A_1) [of (A), of course]. The ${}^{*}_{\beta}c_{\zeta}^{h,w}(t, \tau)$ are entire in τ. Accordingly,

$$(21) \qquad {}_{\beta}^{*}c_{\zeta}^{h,w}(t, \tau) = \sum_{s=0}^{\infty} \sum_{q=0}^{\infty} {}_{\beta}^{*}c_{\zeta:s,q}^{h,w}\, t^{s} \tau^{q},$$

$$(21a) \qquad {}^{*}g^{h,w,j}(t) = \sum_{s=0}^{\infty} {}^{*}g_{:s}^{h,w,j}\, t^{s};$$

here the involved series converge for $|t| < \rho'$. On using (18), (15a) and (12), (12a) we obtain

$$(22) \qquad {}_{\beta}^{*}c_{\zeta:s,q}^{h,1} = 0 \begin{cases} s + q \leqq \phi - 2 & (\zeta = 1); \\ s + q \leqq \phi - 1 & (\zeta > 1). \end{cases}$$

Assume the truth of the following (for $\sigma = 1$ already established in (22)):

$$(22a) \qquad {}_{\beta}^{*}c_{\zeta:s,q}^{h,\sigma} = 0 \begin{cases} s + q \leqq \phi - 2 & (\zeta \leqq \sigma), \\ s + q \leqq \phi - 1 & (\zeta > 1). \end{cases}$$

In view of (20), (15), (15a), (12) and (12a), the relations (22a) would imply that

$$(23)\qquad {}_{\beta}c^{*\,h,w}_{\zeta:\,s,q} = 0\begin{cases} s+q \leqq \phi - 2 & (\zeta \leqq w), \\ s+q \leqq \phi - 1 & (\zeta > w). \end{cases}$$

By induction it is inferred at once that (23) holds for all values of the involved subscripts and superscripts.

Let P_0 denote P (cf. definition following (15a)) *with a sufficiently small vicinity of the origin excluded.* By (18), (18a), (16), (14) and (14a)

$$(24)\qquad \left| {}_{\beta}c^{*\,h,1}_{\zeta}(t,\tau) \right| < dR'' \left| t \right|^{\phi} \left| \frac{t}{\tau} \right|^{-n} e^{\rho|t-\tau|},$$

$$(24a)\qquad \left| {}^{*}g^{h,1,j}(t) \right| < dR'' \left| t \right|^{\phi} \left| t \right|^{-n} e^{\rho|t|} \qquad (\left| \tau \right| \geqq \bar{p} > 0;\ \bar{p} \text{ small})$$

for t in P_0. It will be proved that, for t in P_0 and R_0 sufficiently great,

$$(25)\qquad \left| {}_{\beta}c^{*\,h,w}_{\zeta}(t,\tau) \right| < R_0 \left| t \right|^{\phi} \left| \frac{t}{\tau} \right|^{-n} e^{\rho|t-\tau|} \qquad (\left| \tau \right| \geqq \bar{p}),$$

$$(25a)\qquad \left| {}^{*}g^{h,w,j}(t) \right| < R_0 \left| t \right|^{\phi} \left| t \right|^{-n} e^{\rho|t|}$$

for all values of the involved subscripts and superscripts. Suppose that, for some positive R_1, for t in P_0 and $|\tau| \geqq \bar{p}$,

$$(26)\qquad \left| {}_{\beta}c^{*\,h,\sigma}_{\zeta}(t,\tau) \right| < R_1 \left| t \right|^{\phi} \left| \frac{t}{\tau} \right|^{-n} e^{\rho|t-\tau|},$$

$$(26a)\qquad \left| {}^{*}g^{h,\sigma,j}(t) \right| < R_1 \left| t \right|^{\phi} \left| t \right|^{-n} e^{\rho|t|}$$

$$(\sigma = 1, \cdots, w-1;\ 2 \leqq w \leqq p).$$

By (20), (16), (14), (16a) we would have in virtue of (26)

$$(27)\qquad \left| {}_{\beta}c^{*\,h,w}_{\zeta}(t,\tau) \right| < \left| t \right|^{\phi} \left| \frac{t}{\tau} \right|^{-n} e^{\rho|t-\tau|} d(R'' + dpR_1).$$

On the other hand, in view of (20a), (16), (14a) and (16a) the implication of (26a) would be

$$(27a)\qquad \left| {}^{*}g^{h,w,j}(t) \right| < \left| t \right|^{\phi} \left| t \right|^{-n} e^{\rho|t|}.$$

Here (27) and (27a) have been established for t in P_0 and for $|\tau| \geqq \bar{p}$. The truth of (25) and (25a) follows by induction.

We formulate the above developments of this section in the Lemma.

LEMMA 3. *The formal series* ${}_{h}\bar{\eta}_{w}^{j-1}(t)$ (*cf.* (3a)), *connected with the formal series* (1; §4) *by means of* (1a; §4), *satisfy a certain integral system* (A_3). *The coefficients of this system are defined by convergent series* (21), (21a); *they are meromorphic functions in t, whose only finite t-singularities are poles at the non-zero roots of the characteristic equation of* (A_1) (*of* (A), *also*). *The* ${}_{\beta}\overset{*}{c}{}_{\zeta}^{h,w}(t,\tau)$ *are entire in* τ. *Moreover, inasmuch as properties at* ($t=0, \tau=0$) *are concerned, these coefficients satisfy* (23). *On the other hand, essential properties at infinity are characterized by* (25) *and* (25a); *these inequalities are valid for* $|\tau| \geqq \bar{\rho} > 0$ ($\bar{\rho}$ *sufficiently small*) *and for t in* P_0 (*cf. definition preceding* (24)).

6. **The dominant system of integral equations.** In proving convergence of the formal solutions ${}_{h}\bar{\eta}_{w}^{j-1}(t)$ (3a; §5), of the integral system (A_3), the method of successive approximations (used with success by J. Horn in analogous, but simpler, situations) leads to apparently unsurmountable algebraic difficulties. Accordingly, a different method will be employed. We shall establish a dominant system of integral equations; that is, a system from the convergence of whose solutions convergence of a set of solutions of (A_3) can be inferred. It will be necessary, first, to derive in detail the relations satisfied by the coefficients of the ${}_{h}\bar{\eta}_{w}^{j-1}(t)$. On using (3a; §5) and (21; §5) it follows that

$$\int_0^t {}_{\beta}\overset{*}{c}{}_{\zeta}^{h,w}(t,\tau)\, {}_{h+\beta}\bar{\eta}_{\zeta}^{j-1}(\tau)\,d\tau = \sum_{s=0}^{\infty}\sum_{H=0}^{\infty}\sum_{\nu=0}^{\infty} {}_{\gamma}\overset{*}{c}{}_{\zeta:s,H}^{h,w}\, {}_{h+\beta}\bar{\eta}_{\zeta:\nu}^{j-1} \frac{t^{s+H+\nu+1}}{H+\nu+1}. \tag{1}$$

Also, by (3a; §5),

$$t^{\phi}\, {}_{h}\bar{\eta}_{w}^{j-1}(t) = \sum_{q=\phi-1}^{\infty} {}_{h}\bar{\eta}_{w:\,q-\phi+1}^{j-1}\, t^{q+1}. \tag{1a}$$

Substitution of (1), (1a) and (21a; §5) in (A_3) gives, after a suitable arrangement of terms,

$$\begin{aligned} \sum_{q=\phi-1}^{\infty} {}_{h}\bar{\eta}_{w:\,q-\phi+1}^{j-1}\, t^{q+1} &= \sum_{q=0}^{\infty}\left[\sum_{\beta=0}^{j-1-h}\sum_{\zeta=1}^{p}\sum_{\nu=0}^{q}\left(\sum_{H=0}^{q-\nu}\frac{{}_{\beta}\overset{*}{c}{}_{\zeta:\,q-H-\nu,H}^{h,w}}{H+\nu+1}\right){}_{h+\beta}\bar{\eta}_{\zeta:\,\nu}^{j-1}\right]t^{q+1} \\ &\qquad + \sum_{q}^{\infty} \overset{*}{g}{}_{q+1}^{h,w,j}\, t^{q+1} \\ &= \sum_{q} {}_{h}f_{w:\,q}\, t^{q+1} \quad (h=0,1,\cdots,j-1;\ w=1,\cdots,p). \end{aligned} \tag{2}$$

By virtue of (23; §5) it is concluded that in (2)

$$ {}_{h}f_{w:\,q} = 0 \qquad (h=0,1,\cdots,j-1;\ w=1,\cdots,p;\ q \leqq \phi-2).$$

Thus (2) is formally possible if

$$
{}_{h}\bar{\eta}^{j-1}_{w:\,q-(\phi-1)} = \sum_{\beta=0}^{j-h-1} \sum_{\zeta=1}^{p} \sum_{\nu=0}^{q-\phi+1} {}_{\beta}K^{h,w}_{\zeta,\nu:\,q}({}^{*}c)\; {}_{h+\beta}\bar{\eta}^{j-1}_{\zeta:\,\nu} + {}^{*}g^{h,w,j}_{q+1} \tag{3}
$$

$$
(q = \phi - 1, \phi, \phi + 1, \cdots ; h = 0, 1, \cdots , j - 1; w = 1, \cdots , p),
$$

where, for $\zeta > w$, the summation with respect to ν is extended from $\nu = 0$ to $\nu = q - \phi$, and

$$
{}_{\beta}K^{h,w}_{\zeta,\nu:\,q}({}^{*}c) = \sum_{H=0}^{q-\nu} \frac{1}{H + \nu + 1}\; {}_{\beta}^{*}c^{h,w}_{\zeta:\,q-\nu-H,H}. \tag{3a}
$$

From (3) and (3a) it is inferred that

$$
\begin{aligned}
K(h, w, q; {}^{*}c)\; {}_{h}\bar{\eta}^{j-1}_{w:\,q-(\phi-1)} = {}&{}^{*}g^{h,w,j}_{q+1} + \sum_{\beta=1}^{j-1-h} {}_{\beta}K^{h,w}_{w,q-\phi+1:\,q}({}^{*}c)\; {}_{h+\beta}\bar{\eta}^{j-1}_{w:\,q-(\phi-1)} \\
&+ \sum_{\beta=0}^{j-h-1} \sum_{\zeta=1}^{w-1} {}_{\beta}K^{h,w}_{\zeta,q-\phi+1:q}({}^{*}c)\; {}_{h+\beta}\bar{\eta}^{j-1}_{\zeta:q-(\phi-1)} + \sum_{\beta=0}^{j-h-1} \sum_{\zeta=1}^{p} \sum_{\nu=0}^{q-\phi} {}_{\beta}K^{h,w}_{\zeta,\nu:q}({}^{*}c)\; {}_{h+\beta}\bar{\eta}^{j-1}_{\zeta:\nu}
\end{aligned} \tag{4}
$$

$$
(h = 0, 1, \cdots , j - 1; w = 1, \cdots , p; q = \phi - 1, \phi, \phi + 1, \cdots),
$$

where

$$
K(h, w, q; {}^{*}c) = 1 - \sum_{H=0}^{\phi-1} \frac{1}{H + q - \phi + 2}\; {}_{0}^{*}c^{h,w}_{w:\,\phi-1-H,H}. \tag{4a}
$$

We note that the $K(h, w, q; {}^{*}c)$ may vanish, if at all, only for a finite number of values of h, w, q. The equation (4) may be solved in succession according to the scheme

$$
q = \phi - 1 \left\{ \begin{array}{l} h = j - 1\ (w = 1, \cdots , p) \\ h = j - 2\ (w = 1, \cdots , p) \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ h = 0 \qquad (w = 1, \cdots , p) \end{array} \right| ; \quad q = \phi \left\{ \quad \cdots \quad \right| ; \quad \cdots \quad \cdots \quad \cdots . \tag{5}
$$

If some of the $K(h, w, q; {}^{*}c)$ are zero, a remark entirely analogous to that following (24; §3) will be valid, and some of the $\bar{\eta}^{j-1}$ may be arbitrary (this, of course, corresponds to the possible arbitrariness of some of the η^{j-1}). It is also observed that none of the ${}^{*}c$, occurring in (4a), are involved in the second members of (4).

Let $(\bar{4})$ *denote the set of equations* (4), *where the* $\bar{\eta}^{j-1}$, *the* ${}^{*}g^{h,w,j}_{q+1}$ *and the* ${}^{*}c$ *are replaced by* $\bar{\eta}^{j-1}$, $\bar{g}^{h,w,j}_{q+1}$ *and* $\bar{c}$, *respectively*. Choose

$$
{}_{0}\bar{c}^{h,w}_{w:\,\phi-1-H,H} = c > 0 \tag{6}
$$

$$
(H = 0, 1, \cdots , \phi - 1; w = 1, \cdots , p; h = 0, \cdots , j - 1),
$$

where c is sufficiently small so that

(6a) $K(h, w, q; \bar{c}) > 0 \qquad (q \geqq \phi - 1; w = 1, \cdots, p; h = 0, \cdots, j-1).$

By (6a) and since

$$\lim_{q\to\infty} K(h, w, q; {}^*c) = \lim_{q\to\infty} K(h, w, q; \bar{c}) = 1,$$

we have, excluding the values h, w, q for which $K(h, w, q; {}^*c) = 0$,

$$0 < \frac{K(h, w, q; \bar{c})}{| K(h, w, q; {}^*c) |} < B(> 1) \tag{6b}$$

$$(w = 1, \cdots, p; h = 0, \cdots, j-1; q \geqq \phi - 1).$$

Furthermore, take

$${}_{\beta}\overset{h,w}{\bar{c}}_{\zeta:\rho,u} = 0 \begin{cases} \rho + u \leqq \phi - 2 & (1 \leqq \zeta \leqq w), \\ \rho + u \leqq \phi - 1 & (w < \zeta \leqq p); \end{cases} \tag{7}$$

$$\overset{h,w}{\bar{g}}_{q+1} = 0 \qquad (q = -1, 0, \cdots, \phi - 2); \tag{7a}$$

on the other hand, *let the other* $\bar{c}$ *and* $\bar{g}$ (that is, those not referred to in (6), (7), (7a)) *be positive and greater than the absolute values of the corresponding* ${}^*c$ and ${}^*g$. Since $K(h, w, q; \bar{c}) \neq 0$ for any of the involved values of h, w, q it is observed that *equations* $(\bar{4})$ *determine all the* ${}_h\bar{\eta}^{j-1}_{w:l}$ *uniquely.* By virtue of (6a), and in view of the assumptions concerning the $\bar{c}$ and the $\bar{g}$ we conclude that *the* ${}_h\bar{\eta}^{j-1}_{w:l}$ *are all positive.* We now prove that, with B of (6b) sufficiently great,

$$| {}_h\bar{\eta}^{j-1}_{w:l} | < C^{l+1} \, {}_h\bar{\eta}^{j-1}_{w:l} \tag{8}$$

$$(C = B^{jp}; h = 0, 1, \cdots, j-1; w = 1, \cdots, p; l = 0, 1, \cdots).$$

Of course, when there is any arbitrariness in the determination of the $\bar{\eta}^{j-1}$, we shall take some particular set of the $\bar{\eta}^{j-1}$. By (4a) there exists an integer $q_0 \geqq \phi - 1$ such that

$$K(h, w, q; {}^*c) \neq 0 \tag{9}$$

$$(h = 0, 1, \cdots, j-1; w = 1, \cdots, p; q = q_0, q_0 + 1, \cdots).$$

Let the $\bar{\eta}^{j-1}$ be assigned some particular admissible set of values. There will exist then a number B so that not only (6b) holds but also

$$| {}_h\bar{\eta}^{j-1}_{w:l} | < C^{l+1} \, {}_h\bar{\eta}^{j-1}_{w:l} \tag{9a}$$

$$(C = B^{jp}; h = 0, 1, \cdots, j-1; w = 1, \cdots, p; l = 0, 1, \cdots, q_0 - \phi + 1).$$

The essential fact is observed that, for $q \geqq q_0$, equations (4) define the

${}_{h}\bar{\eta}^{j-1}_{w:\,q-\phi+1}$ ($q \geqq q_0$) uniquely in terms of certain other $\bar{\eta}^{j-1}$, previously defined or selected in a definite manner. Suppose now that

$$\left| {}_{h}\bar{\eta}^{j-1}_{w:\,l} \right| < C^{l+1}\, {}_{h}\bar{\eta}^{j-1}_{w:\,l} \tag{9b}$$
$$(h = 0, 1, \cdots, j-1;\ w = 1, \cdots, p;\ l = 0, \cdots, q-\phi)$$

where $q-1 \geqq q_0$ (this secures $q-\phi \geqq q_0-\phi+1$).

Since the ${}_{\beta}K_{\zeta,i:\sigma}(^*c)$ are forms, with positive coefficients, linear in the $^*c$, it follows that

$$\left| {}_{\beta}K^{h,w}_{\zeta,i:\sigma}(^*c) \right| < {}_{\beta}K^{h,w}_{\zeta,i:\sigma}(\bar{c}) \tag{10}$$

for all involved values of the subscripts and superscripts. In view of (10) and (6b), application of (9b) to (4) would yield the inequalities

$$\begin{aligned} \left| {}_{h}\bar{\eta}^{j-1}_{w:\,q-\phi+1} \right| < \frac{B}{K(h, w, q; \bar{c})} \Bigg[& \bar{g}^{h,w,j}_{q+1} + \sum_{\beta=1}^{j-1-h} {}_{\beta}K^{h,w}_{w,q-\phi+1:\,q}(\bar{c}) \left| {}_{h+\beta}\bar{\eta}^{j-1}_{w:\,q-\phi+1} \right| \\ & + \sum_{\beta=0}^{j-h-1} \sum_{\zeta=1}^{w-1} {}_{\beta}K^{h,w}_{\zeta,q-\phi+1:\,q}(\bar{c}) \left| {}_{h+\beta}\bar{\eta}^{j-1}_{\zeta:\,q-\phi+1} \right| \\ & + \sum_{\beta=0}^{j-h-1} \sum_{\zeta=1}^{p} \sum_{\nu=0}^{q-\phi} {}_{\beta}K^{h,w}_{\zeta,\nu:\,q}(\bar{c})\, {}_{h+\beta}\bar{\eta}^{j-1}_{\zeta:\,\nu} C^{\nu+1} \Bigg] \end{aligned} \tag{11}$$

From (11; $h=j-1$; $w=1$)

$$\begin{aligned} \left| {}_{j-1}\bar{\eta}^{j-1}_{1:\,q-\phi+1} \right| & < \frac{B}{K(j-1, 1, q; \bar{c})} \left[\bar{g}^{j-1,1,j}_{q+1} + \sum_{\zeta=1}^{p} \sum_{\nu=0}^{q-\phi} {}_{0}K^{j-1,1}_{\zeta,\nu:\,q}(\bar{c})\, {}_{j-1}\bar{\eta}^{j-1}_{\zeta:\,\nu} C^{\nu+1} \right] \\ & < BC^{q-\phi+1}\, {}_{j-1}\bar{\eta}^{j-1}_{1:\,q-\phi+1}. \end{aligned} \tag{11a}$$

The latter inequality is obtained using the fact that $C>1$ and that the expression within the brackets in the second member of (11a) would be equal to $K(j-1, 1, q; \bar{c})$ if C were replaced by unity. This type of reasoning is employed in the sequel, but no explicit reference will be made to it. Assume that

$$\left| {}_{j-1}\bar{\eta}^{j-1}_{\zeta:\,q-\phi+1} \right| < B^{\zeta} C^{q-\phi+1}\, {}_{j-1}\bar{\eta}^{j-1}_{\zeta:\,q-\phi+1} \quad (\zeta = 1, \cdots, w-1;\ 2 \leqq w \leqq p). \tag{11b}$$

By virtue of (11b), (11) would imply that

$$\begin{aligned} \left| {}_{j-1}\bar{\eta}^{j-1}_{w:\,q-\phi+1} \right| < \frac{B^{w} C^{q-\phi+1}}{K(j-1, w, q; \bar{c})} \Bigg[& \bar{g}^{j-1,w}_{q+1} B^{-(w-1)} C^{-(q-\phi+1)} \\ & + \sum_{\zeta=1}^{w-1} {}_{0}K^{j-1,w}_{\zeta,q-\phi+1:\,q}(\bar{c})\, {}_{j-1}\bar{\eta}^{j-1}_{\zeta:\,q-\phi+1} B^{-(w-1-\zeta)} \\ & + \sum_{\zeta=1}^{p} \sum_{\nu=0}^{q-\phi} {}_{0}K^{j-1,w}_{\zeta,\nu:\,q}(\bar{c})\, {}_{j-1}\bar{\eta}^{j-1}_{\zeta:\,\nu} C^{-(q-\phi-\nu)} B^{-(w-1)} \Bigg] \\ < B^{w} C^{q-\phi+1}\, {}_{j-1}\bar{\eta}^{j-1}_{w:\,q-\phi+1}. & \end{aligned} \tag{12}$$

By induction it is seen that (11b) is valid for $\zeta=1, \cdots, p$. In particular

$$(13) \qquad \left| {}_{j-1}\bar{\eta}^{j-1}_{\zeta:\,q-\phi+1} \right| < B^{p} C^{q-\phi+1}\, {}_{j-1}\bar{\eta}^{j-1}_{\zeta:\,q-\phi+1} \qquad (\zeta = 1, \cdots, p).$$

Suppose that, for $\lambda=j-1, j-2, \cdots, h+1$ $(0 \leqq h \leqq j-2)$,

$$(14) \qquad \left| {}_{\lambda}\bar{\eta}^{j-1}_{\zeta:\,q-\phi+1} \right| < B^{(j-\lambda)p} C^{q-\phi+1}\, {}_{\lambda}\bar{\eta}^{j-1}_{\zeta:\,q-\phi+1} \qquad (\zeta = 1, \cdots, p).$$

By (11) we then would have

$$(14a) \qquad \begin{aligned} \left| {}_{h}\bar{\eta}^{j-1}_{1:\,q-\phi+1} \right| &< \frac{B}{K(h, 1, q; \bar{c})} \Bigg[\bar{g}^{h,1,j}_{q+1} + \sum_{\beta=1}^{j-1-h} {}_{\beta}K^{h,1}_{1,\,q-\phi+1:\,q}(\bar{c})\; {}_{h+\beta}\bar{\eta}^{j-1}_{1:\,q-\phi+1} \\ &\qquad \times C^{q-\phi+1} B^{(j-h-\beta)p} + \sum_{\beta=0}^{j-h-1} \sum_{\zeta=1}^{p} \sum_{\nu=0}^{q-\phi} {}_{\beta}K^{h,1}_{\zeta,\nu:\,q}(\bar{c})\; {}_{h+\beta}\bar{\eta}^{j-1}_{\zeta:\,\nu} C^{\nu+1} \Bigg] \\ &< B B^{(j-h-1)p} C^{q-\phi+1}\, {}_{h}\bar{\eta}^{j-1}_{1:\,q-\phi+1}. \end{aligned}$$

Assume now that

$$(15) \qquad \left| {}_{h}\bar{\eta}^{j-1}_{\zeta:\,q-\phi+1} \right| < B^{(j-h-1)p} B^{\zeta} C^{q-\phi+1}\, {}_{h}\bar{\eta}^{j-1}_{1:\,q-\phi+1}$$
$$(\zeta = 1, \cdots, w-1;\ 2 \leqq w \leqq p);$$

in (14a) these inequalities have been demonstrated for $w=2$. From (11) we would have

$$(15a) \qquad \begin{aligned} \left| {}_{h}\bar{\eta}^{j-1}_{w:\,q-\phi+1} \right| &< \frac{B}{K(h, w, q; \bar{c})} \Bigg[\bar{g}^{h,w,j}_{q+1} + \sum_{\beta=1}^{j-1-h} {}_{\beta}K^{h,w}_{w,\,q-\phi+1:\,q}(\bar{c}) \\ &\qquad \times {}_{h+\beta}\bar{\eta}^{j-1}_{w:\,q-\phi+1} B^{(j-h-\beta)p} C^{q-\phi+1} + \sum_{\beta=1}^{j-h-1} \sum_{\zeta=1}^{w-1} {}_{\beta}K^{h,w}_{\zeta,\,q-\phi+1:\,q}(\bar{c})\; {}_{h+\beta}\bar{\eta}^{j-1}_{\zeta:\,q-\phi+1} \\ &\qquad \times B^{(j-h-\beta)p} C^{q-\phi+1} + \sum_{\zeta=1}^{w-1} {}_{0}K^{h,w}_{\zeta,\,q-\phi+1:\,q}(\bar{c})\; {}_{h}\bar{\eta}^{j-1}_{\zeta:\,q-\phi+1} B^{(j-h-1)p} \\ &\qquad \times B^{\zeta} C^{q-\phi+1} + \sum_{\beta=0}^{j-h-1} \sum_{\zeta=1}^{p} \sum_{\nu=0}^{q-\phi} {}_{\beta}K^{h,w}_{\zeta,\nu:\,q}(\bar{c})\; {}_{h+\beta}\bar{\eta}^{j-1}_{\zeta:\,\nu} C^{\nu+1} \Bigg] \\ &< B^{(j-h-1)p} B^{w} C^{q-\phi+1}\, {}_{h}\bar{\eta}^{j-1}_{w:\,q-\phi+1}. \end{aligned}$$

Thus (15) holds for $\zeta=w$ and, consequently, for $\zeta=1, \cdots, p$. Whence it follows that

$$\left| {}_{h}\bar{\eta}^{j-1}_{\zeta:\,q-\phi+1} \right| < B^{(j-h)p} C^{q-\phi+1}\, {}_{h}\bar{\eta}^{j-1}_{\zeta:\,q-\phi+1} \qquad (\zeta = 1, \cdots, p).$$

This, however, means that (14) would hold for $\lambda=h$ if (14) holds for $\lambda=j-1$, $j-2, \cdots, h+1$. For $\lambda=j-1$ (14) is observed to be true; thus, (14) holds for

$\lambda = j-1, j-2, \cdots, 0$. The latter fact implies that

$$(16) \quad \left| {}_h\bar{\eta}^{j-1}_{\zeta:\, q-\phi+1} \right| < C^{q-\phi+2} \; {}_h\eta^{j-1}_{\zeta:\, q-\phi+1} \qquad (h = 0, 1, \cdots, j-1; \zeta = 1, \cdots, p).$$

Since (16) is a consequence of (9b), inequalities (8) are seen to be proved. We observe that the following fact has been demonstrated.

Let the $\bar{c}$ *and the* $\bar{g}$ *be numbers satisfying* (6), (6a), (*cf.* (4a)), (7), (7a) *and also satisfying the italicized statement immediately following* (7a). *Let the coefficients in the series*

$$(17) \qquad {}_h\bar{\eta}_w(t) = \sum_{\nu=0}^{\infty} {}_h\bar{\eta}^{j-1}_{w:\,\nu} t^{\nu} \qquad (h = 0, \cdots, j-1; w = 1, \cdots, p)$$

be defined by the equations $(\bar{4})$, *formed with the mentioned* $\bar{c}$ *and* $\bar{g}$ *and characterized in the italics preceding* (6). *In view of* (8), *whenever the series* (17) *all converge in a vicinity of* $t=0$, *the same will be true* (*in a possibly smaller vicinity of* $t=0$) *for any set of formal solutions* ${}_h\bar{\eta}_w(t)$ (*cf.* (3a; §5)) *of the integral system* (A_3).

We are now in the position to form a system of integral equations

$$(\bar{A}_3) \qquad t^{\phi} \, {}_h\bar{\eta}^{j-1}_w(t) = \sum_{\beta=0}^{j-h-1} \sum_{\zeta=1}^{p} \int_0^t {}_\beta\bar{c}^{h,w}_{\zeta}(t, \tau) \; {}_{h+\beta}\bar{\eta}^{j-1}_{\zeta}(\tau) d\tau + \bar{g}^{h,w,j}(t)$$

$$(h = 0, \cdots, j-1; w = 1, \cdots, p);$$

here

$$(18) \qquad {}_\beta\bar{c}^{h,w}_{\zeta}(t, \tau) = \sum_{s=0}^{\infty} \sum_{q=0}^{\infty} {}_\beta\bar{c}^{h,w}_{\zeta:\,s,q} t^{s} \tau^{q}, \qquad \bar{g}^{h,w,j}(t) = \sum_{s=0}^{\infty} \bar{g}^{h,w,j}_s t^{s}.$$

This system is satisfied by the set of formal series (17), defined by $(\bar{4})$. Choose c' sufficiently small. Let r be a sufficiently large positive number. Moreover, we shall choose

$$(18a) \qquad {}_\beta\bar{c}^{h,w}_{\zeta}(t, \tau) = \sum_{i=\phi-1}^{\infty} r^i \sum_{H=0}^{i} t^{i-H} \tau^{H} \qquad (\zeta \leq w),$$

$$(18b) \qquad {}_\beta\bar{c}^{h,w}_{\zeta}(t, \tau) = \sum_{i=\phi}^{\infty} r^i \sum_{H=0}^{i} t^{i-H} \tau^{H} \qquad (\zeta > w),$$

$$(18c) \qquad \bar{g}^{h,w,j}(t) = \sum_{i=\phi}^{\infty} r^{i-\phi+1} t^{i},$$

except that

$$(18d) \qquad {}_0\bar{c}^{h,w}_{w}(t, \tau) = c' r^{\phi-1} \sum_{H=0}^{\phi-1} t^{\phi-1-H} \tau^{H} + \sum_{i=\phi}^{\infty} r^i \sum_{H=0}^{i} t^{i-H} \tau^{H}.$$

These series, of course, all converge in the vicinity of $(t=0, \tau=0)$. It is seen that the constants, occurring in the second members of $(\bar{4})$, will then be

$$
{}_{\gamma}\bar{c}^{h,w}_{\zeta:\rho,u} = r^{\rho+u}, \qquad \bar{g}^{h,w}_{q+1} = r^{q-\phi+2}, \tag{19}
$$

while

$$
{}_{0}\bar{c}^{h,w}_{w:\phi-1-H,H} = c' r^{\phi-1} = \bar{c}. \tag{19a}
$$

The corresponding set of equations $(\bar{4})$ is not practicable for a proof of convergence of the solutions (17) of $(\bar{A}_3)$. Hence use will be made of a different set of relations, obtainable by utilizing the special form of $(\bar{A}_3)$. In fact, (18a), (18b), (18c), (18d) are equivalent to

$$
{}_{\beta}\bar{c}^{h,w}_{\zeta}(t, \tau) = \frac{1}{(1 - rt)(1 - r\tau)} - \sum_{\rho+u \leqq \phi-2} r^{\rho+u} t^{\rho} \tau^{u} \qquad (\zeta \leqq w), \tag{20}
$$

$$
{}_{\beta}\bar{c}^{h,w}_{\zeta}(t, \tau) = \frac{1}{(1 - rt)(1 - r\tau)} - \sum_{\rho+u \leqq \phi-1} r^{\rho+u} t^{\rho} \tau^{u} \qquad (\zeta > w), \tag{20a}
$$

$$
\bar{g}^{h,w,i}(t) = rt^{\phi}/(1 - rt), \tag{20b}
$$

except that

$$
{}_{0}\bar{c}^{h,w}_{w}(t, \tau) = c' r^{\phi-1} \sum_{H=0}^{\phi-1} t^{\phi-1-H} \tau^{H} + \frac{1}{(1 - rt)(1 - r\tau)} - \sum_{\rho+u \leqq \phi-1} r^{\rho+u} t^{\rho} \tau^{u}. \tag{20c}
$$

Substitution of (20), (20a), (20b), (20c) in $(\bar{A}_3)$ results in

$$
\begin{aligned}
(t^{\phi} - rt^{\phi+1})\, {}_{h}\bar{\eta}^{j-1}_{w}(t) = {} & \sum_{\beta=0}^{j-1-h} \sum_{\zeta=1}^{p} \int_0^t \frac{{}_{h+\beta}\bar{\eta}^{j-1}_{\zeta}(\tau)\,d\tau}{1 - r\tau} \\
& - (1 - c) r^{\phi-1} \sum_{H=0}^{\phi-1} (t^{\phi-1-H} - rt^{\phi-H}) \int_0^t \tau^{H}\, {}_{h}\bar{\eta}^{j-1}_{w}(\tau)\,d\tau \\
& - \sum_{\beta=0}^{j-1-h} \sum_{\zeta=1}^{p} \sum_{\rho+u \leqq \phi-2} (r^{\rho+u} t^{\rho} - r^{\rho+u+1} t^{\rho+1}) \int_0^t \tau^{u}\, {}_{h+\beta}\bar{\eta}^{j-1}_{\zeta}(\tau)\,d\tau \\
& - r^{\phi-1} \sum_{\beta=0}^{j-1-h} \sum_{\zeta=w+1}^{p} \sum_{u=0}^{\phi-1} (t^{\phi-1-u} - rt^{\phi-u}) \int_0^t \tau^{u}\, {}_{h+\beta}\bar{\eta}^{j-1}_{\zeta}(\tau)\,d\tau + rt^{\phi}
\end{aligned} \tag{21}
$$

$$
(h = 0, \cdots, j - 1;\ w = 1, \cdots, p).
$$

On using (17) the following formal relations are obtained:

$$
\int_0^t \frac{{}_{h+\beta}\bar{\eta}^{j-1}_{\zeta}(\tau)\,d\tau}{1 - r\tau} = \sum_{s=-\phi+1}^{\infty} \left(\sum_{\lambda=0}^{s+\phi-1} {}_{h+\beta}\bar{\eta}^{j-1}_{\zeta:\lambda} \frac{r^{s+\phi-1-\lambda}}{\phi + s} \right) t^{s+\phi}, \tag{21a}
$$

$$(21\text{b})\qquad \int_0^t \tau^u\, {}_{h+\beta}\bar{\eta}_{\zeta}^{j-1}(\tau)d\tau = \sum_{s=0}^{\infty} {}_{h+\beta}\bar{\eta}_{\zeta:s}^{j-1} \frac{t^{s+u+1}}{s+u+1},$$

$$(21\text{c})\qquad (t^{\phi-1-H} - rt^{\phi-H}) \int_0^t \tau^H\, {}_{h}\bar{\eta}_{w}^{j-1}(\tau)d\tau = \sum_{s=0}^{\infty} \left(\frac{{}_{h}\bar{\eta}_{w:s}^{j-1}}{s+H+1} - \frac{r\, {}_{h}\bar{\eta}_{w:s-1}^{j-1}}{s+H} \right) t^{s+\phi};$$

moreover, by virtue of the relationship

$$\sum_{s=0}^{\infty} \sum_{i=0}^{\phi-2} a_{s,i}\, t^{s+i+1} = \sum_{s=-\phi+1}^{\infty} \sum_{i=0}^{\phi-2} a_{\phi+s-1-i,i}\, t^{\phi+s} \qquad (\text{here } a_{\sigma,i} = 0 \text{ for } \sigma < 0),$$

it follows that

$$(21\text{d})\qquad \begin{aligned} &\sum_{\rho+u \leqq \phi-2} (r^{\rho+u} t^{\rho} - r^{\rho+u+1} t^{\rho+1}) \int_0^t \tau^u\, {}_{h+\beta}\bar{\eta}_{\zeta}^{j-1}(\tau)d\tau \\ &= \sum_{s=-\phi+1}^{\infty} \sum_{i=0}^{\phi-2} \sum_{u=0}^{i} \left[\frac{r^i\, {}_{h+\beta}\bar{\eta}_{\zeta:\phi+s-1-i}^{j-1}}{\phi+s-i+u} - \frac{r^{i+1}\, {}_{h+\beta}\bar{\eta}_{\zeta:\phi+s-2-i}^{j-1}}{\phi+s-1-i+u} \right] t^{\phi+s}. \end{aligned}$$

Substituting (21a), (21b), (21c), (21d) in (21) and observing that for $s<0$ the coefficient of $t^{\phi+s}$ in the second member of (21d) may be written as

$$\sum_{\lambda=0}^{s+\phi-1} {}_{h+\beta}\bar{\eta}_{\zeta:\lambda}^{j-1} \frac{r^{s+\phi-1-\lambda}}{\phi+s}$$

and that for $s \geqq 0$ this coefficient may be expressed as

$$\sum_{\lambda=s+1}^{\phi+s-1} {}_{h+\beta}\bar{\eta}_{\zeta:\lambda}^{j-1} \frac{r^{\phi+s-1-\lambda}}{\phi+s} - \sum_{u=0}^{\phi-2} \frac{r^{\phi-1}}{u+s+1}\, {}_{h+\beta}\bar{\eta}_{\zeta:s}^{j-1},$$

we obtain

$$(22)\qquad \begin{aligned} {}_{h}\bar{\eta}_{w:s}^{j-1} = {}& r\, {}_{h}\bar{\eta}_{w:s-1}^{j-1} + \sum_{\beta=0}^{j-1-h} \sum_{\zeta=1}^{p} \left(\sum_{\lambda=0}^{s} {}_{h+\beta}\bar{\eta}_{\zeta:\lambda}^{j-1} \frac{r^{s+\phi-1-\lambda}}{\phi+s} \right. \\ & \left. + \sum_{H=0}^{\phi-2} {}_{h+\beta}\bar{\eta}_{\zeta:s}^{j-1} \frac{r^{\phi-1}}{H+s+1} \right) - (1-c') \sum_{H=0}^{\phi-1} \frac{r^{\phi-1}}{s+H+1}\, {}_{h}\bar{\eta}_{w:s}^{j-1} \\ & + (1-c) \sum_{H=0}^{\phi-1} \frac{r^{\phi}}{s+H}\, {}_{h}\bar{\eta}_{w:s-1}^{j-1} - \sum_{\beta=0}^{j-1-h} \sum_{\zeta=w+1}^{p} \sum_{H=0}^{\phi-1} \frac{r^{\phi-1}}{s+H+1}\, {}_{h+\beta}\bar{\eta}_{\zeta:s}^{j-1} \\ & + \sum_{\beta=0}^{j-1-h} \sum_{\zeta=w+1}^{p} \sum_{H=0}^{\phi-1} \frac{r^{\phi}}{s+H}\, {}_{h+\beta}\bar{\eta}_{\zeta:s-1}^{j-1} + \xi(s)r \\ & (s = 0, 1, \cdots; h = 1, \cdots, j-1; w = 1, \cdots, p), \end{aligned}$$

where

$$\xi(s) = \begin{cases} 0 & (s > 0), \\ 1 & (s = 0). \end{cases} \tag{22a}$$

On writing

$$q_s = \sum_{H=0}^{\phi-1} \frac{r^{\phi-1}}{s+H+1}, \qquad \bar{q}_s = \sum_{H=0}^{\phi} \frac{r^{\phi}}{s+H}, \qquad g(s) = 1 - c'q_s, \tag{23}$$

c' being small enough so that $g(s) > 0$ for all $s > 0$, and observing that in (22) all the ${}_h\bar{\eta}^{j-1}_{\zeta:s}$ $(\zeta > w)$ cancel, we have

$$\begin{aligned} g(s)\ {}_h\bar{\eta}^{j-1}_{w:s} = {} & q_s \sum_{\zeta=1}^{w-1} {}_h\bar{\eta}^{j-1}_{\zeta:s} + q_s \sum_{\beta=1}^{j-1-h} \sum_{\zeta=1}^{w} {}_{h+\beta}\bar{\eta}^{j-1}_{\zeta:s} \\ & + (rg(s-1) + \bar{q}_{s-1})\ {}_h\bar{\eta}^{j-1}_{w:s-1} + \sum_{\beta=0}^{j-1-h} \sum_{\zeta=1}^{w} \frac{r^{\phi}}{\phi+s}\, {}_{h+\beta}\bar{\eta}^{j-1}_{\zeta:s-1} \\ & + \bar{q}_s \sum_{\beta=0}^{j-1-h} \sum_{\zeta=w+1}^{p} {}_{h+\beta}\bar{\eta}^{j-1}_{\zeta:s-1} + \sum_{\beta=0}^{j-1-h} \sum_{\zeta=1}^{p} \sum_{\lambda=0}^{s-2} {}_{h+\beta}\bar{\eta}^{j-1}_{\zeta:\lambda} \frac{r^{s-1-\lambda}r^{\phi}}{\phi+s} \\ & + \xi(s)r \quad (s = 0, 1, \cdots; h = 0, \cdots, j-1; w = 1, \cdots, p). \end{aligned} \tag{24}$$

In view of (23) there exists a number k so that

$$\frac{1}{g(s)}, \qquad (rg(s-1) + \bar{q}_{s-1}) < k, \tag{25}$$

and such that, for $s \geqq 1$,

$$q_s,\ \bar{q}_s,\ \frac{r^{\phi}}{\phi+s} < \frac{k}{s}\,. \tag{25a}$$

Accordingly, the following inequalities are obtained from (24):

$$\begin{aligned} {}_h\bar{\eta}^{j-1}_{w:s} < {} & k^2\left(\sum_{\zeta=1}^{w-1} {}_h\bar{\eta}^{j-1}_{\zeta:s} + \sum_{\beta=1}^{j-1-h} \sum_{\zeta=1}^{w} {}_{h+\beta}\bar{\eta}^{j-1}_{\zeta:s} \right) \\ & + k^2 \sum_{\beta=0}^{j-1-h} \sum_{\zeta=1}^{p} {}_{h+\beta}\bar{\eta}^{j-1}_{\zeta:s-1} + \frac{k^2}{s} \sum_{\beta=0}^{j-1-h} \sum_{\zeta=1}^{p} \sum_{\lambda=0}^{s-2} r^{s-1-\lambda}\ {}_{h+\beta}\bar{\eta}^{j-1}_{\zeta:\lambda} \\ & (h = 0, \cdots, j-1; w = 1, \cdots, p; s = 1, 2, \cdots). \end{aligned} \tag{26}$$

With the aid of (26) convergence of the series (17) will be proved. For some positive a

$${}_h\bar{\eta}^{j-1}_{w:0} < a \qquad (h = 0, 1, \cdots, j-1; w = 1, \cdots, p). \tag{27}$$

Let s_0 $(\geqq 1)$ be a fixed integer. Depending on s_0, there exists a positive num-

ber ρ, which will be taken $>r$, so that

$$ {}_{h}\bar{\eta}^{j-1}_{w:\,l} < a\rho^{l} \tag{27a}$$

$$(l = 0, 1, \cdots, s_0 - 1;\ h = 0, 1, \cdots, j-1;\ w = 1, \cdots, p).$$

By (26), formed for $s=s_0$, it will follow that

$$ {}_{h}\bar{\eta}^{j-1}_{w:\,s_0} < k^2 \sum_{\zeta=1}^{w-1} {}_{h}\bar{\eta}^{j-1}_{\zeta:\,s_0} + k^2 \sum_{\beta=1}^{j-1-h} \sum_{\zeta=1}^{w} {}_{h+\beta}\bar{\eta}^{j-1}_{\zeta:\,s_0} + a\rho_1\rho^{s_0-1} \tag{28}$$

$$(h = 0, 1, \cdots, j-1;\ w = 1, \cdots, p),$$

where $\rho_1 = 2k^2\phi p$ and so is independent of ρ and s_0. In particular, from (28) we have

$$ {}_{j-1}\bar{\eta}^{j-1}_{1:\,s_0} < a\rho_1\rho^{s_0-1}.$$

Suppose now that

$$ {}_{j-1}\bar{\eta}^{j-1}_{\zeta:\,s_0} < a\rho_{w-1}\rho^{s_0-1} \qquad (\zeta = 1, \cdots, w-1;\ 2 \leqq w \leqq p),$$

where ρ_{w-1} is independent of ρ and s. From (28) it would then follow that

$$ {}_{j-1}\bar{\eta}^{j-1}_{w:\,s_0} < a\rho_{w}\rho^{s_0-1} \qquad (\rho_w = k^2 p\rho_{w-1} + \rho_1).$$

Accordingly, by induction we have

$$ {}_{j-1}\bar{\eta}^{j-1}_{\zeta:\,s_0} < ak_1\rho^{s_0-1} \qquad (\zeta = 1, 2, \cdots, p), \tag{29}$$

where k_1 is independent of ρ and s. Assume the truth of the following inequalities, reducing to (29) for $h=j-2$:

$$ {}_{\lambda}\bar{\eta}^{j-1}_{\zeta:\,s_0} < a\, k_{j-h-1}\, \rho^{s_0-1} \tag{30}$$

$$(\lambda = j-1, j-2, \cdots, h+1;\ \zeta = 1, \cdots, p;\ 0 \leqq h \leqq j-2),$$

where k_{j-h-1} is independent of ρ and s_0. From (28) we then find that

$$ {}_{h}\bar{\eta}^{j-1}_{1:\,s_0} < a\, k_{j-h-1,1}\, \rho^{s_0-1},$$

where $k_{j-h-1,1} = k^2\phi p k_{j-h-1} + \rho_1$ so that $k_{j-h-1,1}$ is independent of ρ, s_0. It is furthermore noted that the inequalities

$$ {}_{h}\bar{\eta}^{j-1}_{\zeta:\,s_0} < ak_{j-h-1,w-1}\,\rho^{s_0-1} \qquad (\zeta = 1, \cdots, w-1;\ 2 \leqq w \leqq p), \tag{31}$$

where $k_{j-h-1,w-1}$ is assumed independent of ρ and s_0, together with (30) would imply, by virtue of (28), that

$$ {}_{h}\bar{\eta}^{j-1}_{w:s_0} < ak_{j-h-1,w}\,\rho^{s_0-1}, \tag{31a} $$

$$ k_{j-h-1,w} = k^2 p k_{j-h-1,w-1} + \phi p k_{j-h-1} + \rho_1. \tag{31b} $$

Thus,

$$ {}_{h}\bar{\eta}^{j-1}_{\zeta:s_0} < ak_{j-h}\,\rho^{s_0-1} \qquad (\zeta = 1, \cdots, p), \tag{32} $$

k_{j-h} being independent of ρ and s_0. The inequalities (30) imply (32); whence we conclude that for some $\bar{k}$, independent of ρ and s_0,

$$ {}_{\lambda}\bar{\eta}^{j-1}_{\zeta:s_0} < a\bar{k}\rho^{s_0-1} \qquad (\lambda = j-1, j-2, \cdots, 0; \zeta = 1, \cdots, p).* \tag{33} $$

Choose the number ρ, first introduced in (27a), *so that $\rho \geqq \bar{k}$.* In view of (33) it is then observed that (27a) holds not only for $(l=0, \cdots, s_0-1; h=0, 1, \cdots, j-1; w=1, \cdots, p)$, as originally stated, but necessarily also for $(l=s_0; h=0, \cdots, j-1; w=1, \cdots, p)$. Since inequalities (28) remain valid, when s_0 is replaced by *any* positive integer s, provided only that

$$ {}_{h}\bar{\eta}^{j-1}_{w:l} < a\rho^{l} \tag{34} $$

$$ (l = 0, 1, \cdots, s-1; h = 0, \cdots, j-1; w = 1, \cdots, p), $$

it follows that (34) is valid for $(l=0, 1, \cdots; h=0, \cdots, j-1; w=1, \cdots, p)$.

By virtue of (34) *the formal solutions* (17) *of the dominant system of integral equations* $(\bar{\mathrm{A}}_3)$ *all converge for* $|t| < 1/\rho$.

7. **The main theorem for differential equations.** On taking account of the two italicized statements, one at the very end of §6 and the other preceding $(\bar{\mathrm{A}}_3)$ (§6), the truth of the following lemma becomes evident.

LEMMA 4. *Suppose that a system of integral equations* (A_3) (*cf.* §5) *is given with the following properties*:

(i) *The coefficients of the system are series of the form* (21; §5), (21a; §5) *convergent for* $|t| < \rho'$, $|\tau| < \rho'$.

(ii) (23; §5) *is satisfied.*

(iii) *The system possesses a full set of formal series solutions* ${}_{h}\bar{\eta}^{j-1}_{w}(t)$ *of the form* (3a; §5).

Under these conditions, the elements of any set of formal solutions, referred to in (iii), *will all necessarily converge in some vicinity of the origin.*

It is to be noted that this lemma makes no reference to properties of the coefficients of (A_3) at $(t = \infty, \tau = \infty)$; it also makes no reference to the manner in which such a system might have originated. In view of Lemma 3, it is

* $\bar{k}$, in general, depends on r.

observed that Lemma 4 is applicable to the particular system (A_3), established in §5. *The formal solutions of the latter system, therefore, all converge in some neighborhood of the origin.* We shall now prove the following Lemma, referring to properties at infinity of the analytic continuations of the analytic functions so defined near $t=0$.

LEMMA 5. *Suppose that a system of integral equations* (A_3) *is given which satisfies the conditions of Lemma* 3. *The elements* ${}_h\bar{\eta}_w^{j-1}(t)$ ($h=0, \cdots, j-1$; $w=1, \cdots, p$) *belonging to any particular set of solutions of* (A_3) (*as follows by Lemma* 4, *they are necessarily analytic at* $t=0$) *have analytic continuations in* P_0 (*cf. italics preceding* (24; §5)) *for which*

$$(1)\qquad |{}_h\bar{\eta}_w^{j-1}(t)| < Ce^{q|t|} \qquad (h = 0, 1, \cdots, j-1;\ w = 1, \cdots, p),$$

along every ray $\bar{t}(=\angle t)$ *in* P_0. *In* (1) C *and* q *are positive, sufficiently great*$_1$ *and are independent of* $\bar{t}$ *and* $|t|$.

The proof of this Lemma will be somewhat similar to that which J. Horn gives in proving certain inequalities analogous to (1). We confine ourselves to some particular set of solutions of (A_3). C will be chosen a fixed value such that

$$(2)\qquad |{}_h\bar{\eta}_{w:0}^{j-1}| < C \qquad (h = 0, 1, \cdots, j-1;\ w = 1, \cdots, p).$$

Then

$$(3)\qquad |{}_h\bar{\eta}_w^{j-1}(t)| < Ce^{q|t|} \qquad (h = 0, \cdots, j-1;\ w = 1, \cdots, p)$$

for $|t| \leqq r^0$ (r^0 some positive number) and for every $q \geqq 0$. It will be supposed that the constant $\bar{p}$, occurring in (25; §5), had been chosen suitably small and so that

$$(3a)\qquad 0 < \bar{p} < r^0.$$

Suppose the Lemma is not true. Then, for some $r' > r^0$, the inequalities

$$(4)\qquad |{}_h\bar{\eta}_w^{j-1}(t)| < Ce^{q|t|} \qquad (h = 0, \cdots, j-1;\ w = 1, \cdots, p)$$

will be valid for $|t| < r'$ along every ray $\bar{t}(=\angle t)$ in P (cf. definition following (15a; §5)), while, for some $h=h'$ and $w=w'$, we shall have

$$(4a)\qquad |{}_{h'}\bar{\eta}_{w'}^{j-1}(t)| = Ce^{q|t|} \qquad (t = t' = r'e^{i\bar{t}'};\ \bar{t}' \text{ in } P).$$

It is clear that (4) continues to hold, with $<$ replaced by $\leqq$, when t is on any ray $\bar{t}$ of P_0, while $|t| \leqq r'$. Let the integrations involved below be along the

ray $\bar{t}'$. From (A_3) (§5), by virtue of (4), (4a) and (25a; §5), it then follows that

$$
(5)\qquad \begin{aligned} &\left| t'^{\phi}\ {}_{h'}\bar{\eta}_{w'}^{j-1}(t) \right| = |t'|^{\phi} C e^{q|t'|} \\ &< \sum_{\beta=0}^{j-h-1} \sum_{\zeta=1}^{p} \int_0^{r'} \left| {}_{\beta}\overset{*}{c}{}_{\zeta}^{h,w}(t', \tau) \right| C e^{q|\tau|} d|\tau| + R_0 |t'|^{\phi} |t'|^{-n} e^{\rho|t'|}. \end{aligned}
$$

It is observed that

$$
\sum_{\beta=0}^{j-h-1} \sum_{\zeta=1}^{p} \int_0^{\bar{\rho}} \left| {}_{\beta}\overset{*}{c}{}_{\zeta}^{h,w}(t', \tau) \right| C e^{q|\tau|} d|\tau| < \bar{l} e^{q\bar{\rho}} ,
$$

where $\bar{l}$ is independent of t' and q. Thus from (5), on using (25; §5), we obtain the inequality

$$
(6)\qquad \begin{aligned} (r')^{\phi} C e^{qr'} &< l e^{q\bar{\rho}} + R_0 (r')^{\phi - n} e^{\rho r'} \\ &\quad + \sum_{\beta=0}^{j-h-1} \sum_{\zeta=1}^{p} \int_{\bar{\rho}}^{r'} R_0 (r')^{\phi} e^{\rho(r' - |\tau|)} C e^{q|\tau|} d|\tau|; \end{aligned}
$$

whence

$$
(6a)\qquad \begin{aligned} 1 &< \frac{1}{C} \bar{l} e^{-q(r' - \bar{\rho})} (r')^{-\phi} + \frac{1}{C} R_0 (r')^{-n} e^{-(q-\rho) r'} \\ &\quad + R_0 p \phi \int_{\bar{\rho}}^{r'} e^{-(q-\rho)(r' - |\tau|)} d|\tau|. \end{aligned}
$$

Take $q > \rho$. It is then concluded that the integral in the second member of (6a) satisfies the inequality

$$
(6b)\qquad \int_{\bar{\rho}}^{r'} < \left[\int_0^{r'} = \int_0^{r'} e^{-(q-\rho)x} dx < \int_0^{\infty} e^{-(q-\rho)x} dx = \right] \frac{1}{q - \rho} \cdot
$$

Since $r' > r_0 > \bar{\rho}$, in view of (6b) from (6a) we obtain

$$
(7)\qquad 1 < \frac{1}{C} \bar{l} e^{-q(r^0 - \bar{\rho})} (r^0)^{-\phi} + \frac{1}{C} R_0 (r^0)^{-n} e^{-(q-\rho) r^0} + \frac{R_0 p \phi}{q - \rho} = f(q).
$$

Now, r^0 is chosen independent of q. It is further noted that $f(q)$ is continuous in q, for $q > \bar{\rho}$; moreover,

$$
\lim_{q \to \infty} f(q) = 0.
$$

Accordingly, $q(>\rho)$ may be taken sufficiently great so that $f(q) \leqq 1$; for any such value of q a contradiction arises to (7). Therefore the Lemma is seen

to be true. The inequalities (1) of this Lemma will hold, for instance, when q is such that $f(q)=1$.

Consider a particular ray $\bar{l}$ in P. The formal series ${}_h\eta_w^{j-1}(x)$(cf. (3; §5)), even if divergent, give rise to certain analytic functions

$$
{}_h\eta_w^{j-1}(x) = \int_0^\infty {}_h\bar{\eta}_w^{j-1}(t)e^{tx}\,dt \qquad (h = 0, \cdots, j-1;\ w = 1, \cdots, p). \tag{8}
$$

The truth of this statement is a consequence of the following considerations. Suppose x is in a half plane $H[\bar{l}]$, characterised by an inequality of the form

$$
R(e^{i\bar{l}}x) = |\,x\,| \cos(\bar{l} + \bar{x}) < -q'(<0) \qquad (\bar{x} = \angle x); \tag{9}
$$

for the present we shall take $q'=q+\epsilon$ ($\epsilon>0$ and arbitrarily small). It then follows that the integrals (8) are all absolutely convergent. In fact, by (1) and (9),

$$
\begin{aligned}
|\,{}_h\eta_w^{j-1}(x)\,| &\leqq \int_0^\infty |\,{}_h\bar{\eta}_w^{j-1}(t)\,| \exp(|\,t\,|\,|\,x\,| \cos(\bar{l}+\bar{x}))d\,|\,t\,| \\
&< C\int_0^\infty e^{-\epsilon|t|}d\,|\,t\,| = \frac{C}{\epsilon}
\end{aligned}
$$

$$
(x \text{ in } H[\bar{l}];\ h = 0, \cdots, j-1;\ w = 1, \cdots, p).
$$

The conditions (1; §5), (5; §5) have to be satisfied. The first of these is seen to hold for x in $H[\bar{l}]$ since, by virtue of (9),

$$
|\,e^{tx}t^\alpha\,| < e^{-(q+\epsilon)|t|}\,|\,t\,|^\alpha.
$$

On the other hand, the following inequalities will hold for the functions within the brackets of (5; §5):

$$
\begin{aligned}
\left|\, e^{tx}\int_0^t \right| &< Ce^{-(q+\epsilon)|t|}\int_0^{|t|} |\,\tau\,|^k e^{q|\tau|}d\,|\,\tau\,|^{(H)} \\
&< Ce^{-\epsilon|t|}\int_0^{|t|} |\,\tau\,|^k d\,|\,\tau\,|^{(H)} = \frac{Ce^{-\epsilon|t|}\,|\,t\,|^{k+H}}{(k+1)\cdots(k+H)}
\end{aligned} \tag{9a}
$$

$$
(H = 1, 2, \cdots;\ k = 0, \cdots, n).
$$

Since the last expression above vanishes for $t=0$ and approaches zero as $|t|\to\infty$, the truth of (5; §5) becomes evident.

Thus, the conditions of certain theorems, due to N. E. Nörlund and of central importance in the theory of factorial series,* are seen to be valid, a fact made manifest after certain obvious transformations are carried out. We thus obtain, whenever the ray $\angle t = \bar{l}$ is in P,

* N. E. Nörlund, *Leçons sur les Séries d'Interpolation*, Paris, 1926; pp. 206–208. Also see his theorem on p. 203.

$$(10)\qquad {}_h\eta_w^{j-1}(x) = \sum_{s=0}^{\infty} \frac{{}_h a_{w:s}^{j-1}(\bar{l})}{x(x-\gamma)\cdots(x-s\gamma)}$$

$$(h = 0, \cdots, j-1;\ w = 1, \cdots, p;\ \angle\gamma = \bar{\gamma})$$

where $|\gamma|$ is suitably great, $\bar{\gamma} = -\bar{l}$, and the involved series all converge in a certain half plane $H[\bar{l}]$. *These functions are analytic solutions of the system of "mixed" differential equations* (A_2) (§4). It is clear that results of stated type hold for $j=1, \cdots, \phi$. Accordingly, it is observed that the ϕ formal series solutions of (A_1), $s_j(x)$ (cf. (1; §3), (1b; §3)), give rise to a corresponding set of ϕ analytic solutions $y_j(x)$, obtained by replacing the ${}_h\eta_w^{j-1}(x)$ of (1; §3) by the corresponding (convergent) series (10). We are thus ready to formulate the Main Theorem for differential equations.

THEOREM I. *Let a differential equation* (A) (§1) *be given. Suppose that corresponding to a root* ρ_1, *of multiplicity* ϕ, *of the associated characteristic equation there exists a linearly independent set of* ϕ *formal series solutions, all of normal type* (*cf.* §2) *and all forming one logarithmic group. Bring* (A) *to the corresponding form* (A_1) (§2) *and let* $E(\rho)=0$ *be the characteristic equation of* (A_1). *For every* $\bar{l}$, *not coincident with a value of an angle of a non-zero root of* $E(\rho)=0$,* *the following is true.*

(A_1) *possesses a set of* ϕ *linearly independent analytic solutions*

$$(11)\ y_j(x) = e^{Q(x)}x^{r}\sum_{h=0}^{j-1}\log^h x\left({}_h\eta_0^{j-1} + \sum_{w=1}^{p} x^{(p-w)/p}\sum_{s=0}^{\infty}\frac{{}_h a_{w:s}^{j-1}(\bar{l})}{x(x-\gamma)\cdots(x-s\gamma)}\right)$$

$$(j = 1, \cdots, \phi;\quad \angle\gamma = \bar{\gamma} = -\bar{l};\quad Q(x)\ \textit{a polynomial in}\ x^{1/p}),$$

where $|\gamma|$ *is suitably great and the involved series all converge in a certain half plane* $H[\bar{l}]$ (*cf.* (9); q' *in* (9) *sufficiently great*).

The implications of this theorem for (A) are immediate. The ϕ corresponding solutions of (A) will be of the form (11), where x is replaced by a certain power of x; accordingly, the series involved in these solutions will all converge in certain sectors.†

The theorem is of greatest possible completeness in the sense that even normal formal series solutions of (A) do not always lead to convergent factorial series developments, if corresponding to the multiple root, in question, there

* It is supposed, as may be without any loss of generality, that $\rho=0$ is the root corresponding to the formal solutions in question.

† This is a consequence of known properties of factorial series.

is more than one logarithmic group. Consideration of the following example will demonstrate this statement.

The equation

$$L_3(y) \equiv y^{(3)}(x) + ax^{-1}y^{(1)}(x) + dx^{-3}y(x) = 0 \qquad (a \neq 0,\ d \neq 0) \tag{12}$$

is of the form (A_1); its characteristic equation has a root, $\rho=0$, of multiplicity three. Moreover, it has a formal solution

$$y(x) = \sum_{\nu=0}^{\infty} y_\nu x^{-\nu} \qquad (y_0 = 1). \tag{13}$$

In fact, substituting (13) in (12), we obtain

$$- L_3(y) \equiv \sum_{\lambda=0}^{\infty} f_\lambda x^{-\lambda-3}, \qquad f_\lambda = [\lambda(\lambda+1)(\lambda+2) - d]y_\lambda + a(\lambda+1)y_{\lambda+1}$$

so that the equations $f_\lambda=0$ $(\lambda=0, 1, \cdots)$ are seen to be uniquely solvable for the y_ν. We have

$$y_{\lambda+1} = q(\lambda)y_\lambda, \qquad q(\lambda) = \frac{d}{a(\lambda+1)} - \frac{\lambda(\lambda+2)}{a} \qquad (\lambda = 0, 1, \cdots); \tag{14}$$

whence

$$y_{\lambda+1} = q(0)q(1)\cdots q(\lambda) \qquad (\lambda = 0, 1, \cdots). \tag{14a}$$

It can be shown that the number of logarithmic groups is greater than one.

Suppose that corresponding to the normal solution of (12) there is some convergent factorial series development, of the form

$$1 + \sum_{s=0}^{\infty} \frac{a_s}{x(x-\gamma)\cdots(x-s\gamma)}$$

where γ is some real or complex number. By a theorem of Nörlund referred to previously, this function would be expressible by means of the convergent integral

$$\eta(x) = 1 + \int_0^\infty \bar{y}(t)e^{tx}dt,$$

where the integration is extended along a certain ray and where $\bar{y}(t) = \sum_{\lambda=0}^{\infty}\bar{y}_\lambda t^\lambda$ is analytic at $t=0$. Since $\eta(x)$ satisfies (12) it would necessarily follow that

$$\bar{y}_\lambda = \frac{(-1)^{\lambda+1}y_{\lambda+1}}{\lambda!} \qquad (\lambda = 0, 1, \cdots),$$

the $y_{\lambda+1}$ being defined by (14a). Accordingly,

$$\left| \frac{\bar{y}_\lambda}{\bar{y}_{\lambda-1}} \right| = \left| \frac{q(\lambda)}{\lambda} \right| = \left| \frac{d}{a\lambda(\lambda+1)} - \frac{\lambda+2}{a} \right|$$

so that $|\bar{y}_\lambda/\bar{y}_{\lambda-1}| \to \infty$, as $\lambda \to \infty$. Thus, a contradiction arises to analyticity of $\bar{y}(t)$. Hence there exists no convergent factorial series corresponding to (13).

PART II. LINEAR DIFFERENCE EQUATIONS

8. Some preliminary facts concerning difference equations. As had been demonstrated by Birkhoff, the difference equation (B) (§1) possesses in all cases a linearly independent set of n formal series solutions

$$\text{(1)} \qquad s_i(x) = e^{\mu_i x \log x} e^{Q_i(x)} x^{r_i} \sigma_i(x) \qquad (i = 1, \cdots, n),$$

where

$$\text{(1a)} \qquad Q_i(x) = \sum_{\nu=0}^{k_i-1} \overset{i}{q}_\nu x^{(k_i-\nu)/k_i},$$

$$\text{(1b)} \qquad \sigma_i(x) = \sum_{h=0}^{m_i'} \log^h x \; {}_h\eta^{m_i'}(x),$$

$$\text{(1c)} \qquad {}_h\eta^{m_i'}(x) = \sum_{s=0}^{\infty} {}_h\eta_s^{m_i'} x^{-s/k_i} \qquad (h = 0, \cdots, m_i);$$

here the μ_i are certain rational numbers, the m_i' and k_i are integers ($m_i' \geqq 0$; $k_i = r_i' p$; integer $r_i' \geqq 1$). The formal series can be arranged in logarithmic groups, the exponential factor

$$e^{\mu x \log x} e^{Q(x)} x^r$$

being the same for each member of the same group.* It is convenient to group the μ_i as follows:

$$\text{(2)} \qquad \begin{aligned} m_1 &= \mu_1 = \mu_2 = \cdots = \mu_{\alpha_1} > m_2 = \mu_{\alpha_1+1} = \mu_{\alpha_1+2} = \cdots = \mu_{\alpha_2} > \cdots \\ &> m_\lambda = \mu_{\alpha_{\lambda-1}+1} = \mu_{\alpha_{\lambda-1}+2} = \cdots = \mu_{\alpha_\lambda} (= \mu_n) \qquad (\lambda \geqq 1). \end{aligned}$$

It is to be noted that the formal facts continue to hold when the $d_{i,j}(x)$ are divergent series. To each one of the λ μ-groups, specified in (2), there corresponds a certain characteristic equation.

For the case when the coefficients of (B), except for a few positive integral powers of x, are expressible by convergent factorial series of the type

$$\text{(3)} \qquad \sum_{s=0}^{\infty} \frac{a_s}{x(x-\gamma)\cdots(x-\gamma s)} \qquad (\gamma = 1 \text{ or } -1)$$

* The definition for such groups is analogous to a similar definition in §2.

and, besides, (B) can be brought to the form of an equation of "Fuchsian type," Nörlund obtains full sets of solutions, which are expressible by means of convergent factorial series of the form (3),* thus establishing an analogy to the Fuchsian theory in the field of differential equations.

Nörlund also treats the class of equations (B) where the coefficients can be brought to the form of polynomials in x. His other restrictions on the coefficients amount precisely to the following.

(i) There is only one μ-group.

(ii) All the numbers μ_i are zero.

(iii) There exist no anormal formal series solutions.†

On using Laplace transformations, leading to ordinary linear differential equations, he obtains full sets of solutions expressible with the aid of convergent factorial series (3), where γ may be complex and $|\gamma|$ is sufficiently great.‡

In Horn's work we essentially have an equation (B), whose coefficients contain no fractional powers of x, and which has only one μ-group, all the μ_i being zero. Moreover, he assumes that all the roots of the characteristic equation (necessarily, there will be only one such equation) are distinct. Under these conditions, he obtains solutions with the aid of convergent factorial series (3) ($|\gamma|$ sufficiently great; certain values of $\angle\gamma$ excepted).

Contrary to the restrictions of Nörlund's and Horn's works, in the present paper existence of several μ-groups (cf. (2)) is admitted; moreover, the coefficients of (B) are allowed to contain fractional powers of x.

Consider any particular μ-group. The μ_i of this group will all have the same value, say μ_1. By means of a transformation of the form

$$y(x) = e^{\mu x \log x} z(x)$$

(B) will be brought to the form (B_1). Retaining the notation originally used in (B) (cf. §1) the new equation will be considered as the equation (B), whose coefficients are convergent series

$$(4) \qquad d_{n-k}(x) = \sum_{s=0}^{\infty} d_{n-k,s}\, x^{-s/p} \qquad (k = 0, 1, \cdots, n), §$$

where

---

* Cf. Nörlund on difference equations, loc. cit., pp. 28–56; in these pages he also establishes certain other important results.

† That is, the $Q_i(x)$ and the $\sigma_i(x)$ (cf. (1), (1a), (1b), (1c)) contain no fractional powers of x; on the other hand, logarithms in the formal series may be present.

‡ Cf. Nörlund on difference equations, loc. cit., pp. 56–88.

§ In (4), p will be, in general, different from the corresponding integer in (B).

$$\text{(4a)}\qquad \begin{aligned} &d_{0,0} = d_{1,0} = \cdots = d_{m-1,0} = d_{m+H+1,0} = d_{m+H+2,0} = \cdots = d_{n,0} = 0,\\ &d_{m,0},\quad d_{m+H,0} \neq 0 \qquad\qquad (0 \leqq m;\, m \leqq m + H \leqq n). \end{aligned}$$

If the μ-group under consideration is the one corresponding to a greatest μ_i, the equation (B_1) will have coefficients for which (4a), with $m=0$, will hold.* To the μ-group, under consideration, of (B) there corresponds a certain μ-group of (B_1); the μ_i of the latter group will be all zero. A formal solution (with $\mu=0$) of (B_1) is said to be *normal* when the corresponding polynomial $Q(x)$ (1a) and the corresponding series $\sigma(x)$ ((1b), (1c)) are in integral powers $x^{1/p}$. In the contrary case, that is, when $Q(x)$ and $\sigma(x)$ are actually in integral powers of $x^{1/(r'p)}$ (integer $r'>1$), the formal series of (B_1) is termed *anormal*. A formal solution of (B), with μ not necessarily zero, is normal or anormal according as the corresponding solution (with $\mu=0$) of (B_1) is normal or anormal.

Throughout the text, leading up to the Main Theorem for difference equations (§12), we consider a root ρ' of multiplicity ϕ of the characteristic equation, of (B), associated with a particular μ-group. The only restriction, and as will be shown by an example in §12 a necessary one, will be that associated with this root ρ' there is just one logarithmic group and that all the formal solutions in this group are normal. When, as will be done for convenience, the corresponding equation (B_1) ((4), (4a)) is used, the characteristic equation at hand of (B_1) will be

$$E_1(\rho) = \sum_k d_{n-k,0}\, e^{k\rho} = 0. \tag{5}$$

This equation will possess a root of $\rho=\rho_1$ of multiplicity ϕ.

9. **Conditions for existence of formal solutions of the type specified in §8.** A set of ϕ linearly independent normal formal series solutions (of (B_1)), which by hypothesis corresponds to the root $\rho=\rho_1$ of (5; §8), forms a logarithmic group. These solutions will be written in the form

$$s_j(x) = e^{Q(x)}x^r \sum_{h=0}^{j-1} \log^h x \;{}_h\eta^{j-1}(x) \qquad (j = 1, \cdots, \phi), \tag{1}$$

where

$${}_h\eta^{j-1}(x) = \sum_{s=0}^{\infty} {}_h\eta_s^{j-1} x^{-s/p} \qquad (h = 0, 1, \cdots, j-1). \tag{1a}$$

Since, in (1), $Q(x)$ is in powers of $x^{1/p}$ it follows that the transformation

$$y(x) = e^{Q(x)}x^r\bar{y}(x),$$

* If there exists only one μ-group, $m=0$ and $m+H=n$.

applied to (B_1), does not change the form of (B_1); in particular, the coefficients of the transformed equation will also be in powers of $x^{1/p}$. Thus, without loss of generality, it will be assumed that

$$e^{Q(x)}x^r \equiv 1 \tag{1b}$$

and that not all the

$$ {}_{j-1}\eta_0^{j-1} \qquad (j = 1, \cdots, \phi) \tag{1c}$$

are zero. We shall presently find the conditions satisfied by the coefficients of (B_1), when this difference equation possesses solutions (1), (1a), for which (1b) holds.

Given a series

$$\eta(x) = \sum_{s=0}^{\infty} \eta_s x^{-s/p}, \tag{2}$$

we shall have

$$\begin{aligned} \eta(x + k) &= \sum_{s=0}^{\infty} \eta_s x^{-s/p} \sum_{r=0}^{\infty} C_r^{-s/p} k^r x^{-r} \\ &= \eta_0 + \sum_{r=0}^{\infty}\sum_{\nu=0}^{\infty}\sum_{w=1}^{p} C_r^{-(\nu p+w)/p}\, k^r\, \eta_{\nu p+w}\, x^{-((r+\nu)p+w)/p} \\ &= \eta_0[k] + \sum_{\lambda=0}^{\infty}\sum_{w=1}^{p} \eta_{\lambda p+w}[k] x^{-(\lambda p+w)/p}, \end{aligned} \tag{2a}$$

where $\eta_0(k) = \eta_0$ and

$$\eta_{\lambda p+w}[k] = \sum_{\nu=0}^{\lambda} C_{\lambda-\nu}^{-(\nu p+w)/p} k^{\lambda-\nu} \eta_{\nu p+w} \qquad (\lambda = 0, 1, \cdots; w = 1, \cdots, p). \tag{2b}$$

A certain expression for $\log^h (x+k)$ will also be needed. It is observed that

$$\log (x + k) = \log x + \sum_{\alpha=1}^{\infty} \frac{(-1)^{\alpha+1}}{\alpha} k^\alpha x^{-\alpha}.$$

Moreover, it can be shown without difficulty that

$$\left(\sum_{\alpha=1}^{\infty} \frac{(-1)^{\alpha+1}}{\alpha} k^\alpha x^{-\alpha}\right)^\beta = \sum_{\alpha=0}^{\infty} (-1)^{\alpha+\beta} k^\alpha\, l_{\alpha,\beta}\, x^{-\alpha}, \tag{3}$$

where

$$l_{0,0} = 1, \quad l_{\alpha,0} = 0 \quad (\alpha \geqq 1), \qquad l_{\alpha,\beta} = 0 \quad (\alpha < \beta); \tag{3a}$$

the $l_{\alpha,\beta}$ not mentioned in (3a) are all positive. We then have

$$\log^h(x+k) = \sum_{\beta=0}^{h} \log^{h-\beta} x \, C_\beta^h \sum_{\alpha=0}^{\infty} (-1)^{\alpha+\beta} k^\alpha l_{\alpha,\beta} x^{-\alpha}. \tag{3b}$$

By (2), (2a), (3b), on making use of the identity

$$\sum_{h=0}^{j-1} \sum_{\beta=0}^{h} a_{h-\beta} b_{h,\beta} = \sum_{h=0}^{j-1} a_h \sum_{\beta=0}^{j-h-1} b_{h+\beta,\beta},$$

it follows that, formally,

$$s_j(x+k) = \sum_{h=0}^{j-1} \log^h x \sum_{\beta=0}^{j-h-1} \sum_{\alpha=0}^{\infty} C_\beta^{h+\beta}(-1)^{\alpha+\beta} k^\alpha l_{\alpha,\beta} x^{-\alpha} \sum_{s=0}^{\infty} {}_{h+\beta}\eta_s^{j-1}[k] x^{-s/p}. \tag{4}$$

On substituting (4) in (B_1) and on noting (4; §8), it is observed that

$$L(s_j) = \sum_k d_{n-k}(x) s_j(x+k) = \sum_{h=0}^{j-1} \log^h x \, {}_hf^{j-1}(x), \tag{5}$$

where

$${}_hf^{j-1}(x) = \sum_{\alpha=0}^{\infty} \sum_{s=0}^{\infty} \sum_{\sigma=0}^{\infty} a_{\alpha,s,\sigma}^{j,h} x^{-(\alpha+(s+\sigma)/p)}, \tag{5a}$$

$$a_{\alpha,s,\sigma}^{j,h} = \sum_{\beta=0}^{j-h-1} \sum_{k=0}^{n} a_{\alpha,s,\sigma,\beta,k}^{j,h}, \tag{5b}$$

$$\alpha_{\alpha,s,\sigma,\beta,k}^{j,h} = C_\beta^{h+\beta}(-1)^{\alpha+\beta} l_{\alpha,\beta} \, {}_{h+\beta}\eta_s^{j-1}[k] k^\alpha d_{n-k,\sigma}. \tag{5c}$$

Further examination of ${}_hf^{j-1}(x)$ gives

$${}_hf^{j-1}(x) = \sum_{\alpha=0}^{\infty} \sum_{\tau=0}^{\infty} a_{\alpha,\tau}^{j,h} x^{-(\alpha p+\tau)/p}, \tag{6}$$

$$a_{\alpha,\tau}^{j,h} = \sum_{s=0}^{\tau} a_{\alpha,s,\tau-s}^{j,h}. \tag{6a}$$

It follows next that

$${}_hf^{j-1}(x) = \sum_{\alpha=0}^{\infty} \sum_{u=0}^{\infty} \sum_{v=0}^{p-1} a_{\alpha,up+v}^{j,h} x^{-((\alpha+u)p+v)/p} = \sum_{\rho=0}^{\infty} \sum_{\tau=0}^{p-1} {}_hF_{\rho,\tau}^{j-1} x^{-(\rho p+\tau)/p}, \tag{7}$$

where

$${}_hF_{\rho,\tau}^{j-1} = \sum_{\alpha=0}^{\rho} a_{\alpha,(\rho-\alpha)p+\tau}^{j,h}. \tag{7a}$$

In view of (5) and (7) it is clear that the equations

$$(8)\qquad {}_hF^{j-1}_{\rho,\tau} = 0 \qquad (h = 0, \cdots, j-1;\ \rho = 0, 1, \cdots;\ \tau = 0, 1, \cdots, p-1)$$

have to be satisfied. It will be necessary to get the left members of (8) in considerable detail. By (7a), (6a), (5b) and (5c) it follows that

$$(9)\qquad \begin{aligned} {}_hF^{j-1}_{\rho,\tau} &= \sum_{\alpha=0}^{\rho} \sum_{s=0}^{(\rho-\alpha)p+\tau} \sum_{\beta=0}^{j-h-1} \sum_{k=0}^{n} C_\beta^{h+\beta}(-1)^{\alpha+\beta} l_{\alpha,\beta} k^\alpha \\ &\quad \times {}_{h+\beta}\eta_s^{j-1}[k] d_{n-k,(\rho-\alpha)p+\tau-s} \\ &= {}_hF^{j-1}_{\rho,\tau:1} + {}_hF^{j-1}_{\rho,\tau:2}, \end{aligned}$$

where

$$(9a)\qquad {}_hF^{j-1}_{\rho,\tau:1} = \sum_{\alpha=0}^{\rho} \sum_{\beta=0}^{j-h-1} \sum_{k=0}^{n} C_\beta^{h+\beta}(-1)^{\alpha+\beta} l_{\alpha,\beta}\ {}_{h+\beta}\eta_0^{j-1} k^\alpha d_{n-k,(\rho-\alpha)p+\tau},$$

$$(9b)\qquad \begin{aligned} {}_hF^{j-1}_{\rho,\tau:2} &= \sum_{\alpha=0}^{\rho} \sum_{\lambda=0}^{\rho-\alpha} \sum_{w=1}^{p} \sum_{\beta=0}^{j-h-1} \sum_{k=0}^{n} C_\beta^{h+\beta}(-1)^{\alpha+\beta} l_{\alpha,\beta} k^\alpha \\ &\quad \times {}_{h+\beta}\eta^{j-1}_{\lambda p+w}[k] d_{n-k,(\rho-\alpha-\lambda)p+\tau-w} \qquad (d_{i,j} = 0 \text{ for } j < 0). \end{aligned}$$

For convenience let

$$(10)\qquad \sum_{k=0}^{n} k^\nu d_{n-k,i} = \delta_i^\nu \qquad (\delta_i^\nu = 0 \text{ for } i < 0).$$

On using the relation

$$\sum_{\lambda=0}^{\rho-\alpha} \sum_{\nu=0}^{\lambda} = \sum_{\nu=0}^{\rho-\alpha} \sum_{\lambda=\nu}^{\rho-\alpha},$$

by virtue of (2b) we have from (9b)

$$(11)\qquad \begin{aligned} {}_hF^{j-1}_{\rho,\tau:2} &= \sum_{\beta=0}^{j-h-1} \sum_{w=1}^{p} \sum_{\alpha=0}^{\rho} \sum_{\nu=0}^{\rho-\alpha} \left(\sum_{\lambda=\nu}^{\rho-\alpha} C_\beta^{h+\beta} C_{\lambda-\nu}^{-(\nu p+w)/p} (-1)^{\alpha+\beta} l_{\alpha,\beta} \right. \\ &\quad \left. \times\, \delta^{\lambda-\nu+\alpha}_{(\rho-\alpha-\lambda)p+\tau-w} \right) {}_{h+\beta}\eta^{j-1}_{\nu p+w}. \end{aligned}$$

An application of the relationship

$$\sum_{\alpha=0}^{\rho} \sum_{\nu=0}^{\rho-\alpha} = \sum_{\nu=0}^{\rho} \sum_{\alpha=0}^{\rho-\nu}$$

to (11) will give from (9)

$$(12)\qquad {}_hF^{j-1}_{\rho,\tau} = \sum_{\beta=0}^{j-h-1} {}_hJ^{\rho,\tau}_{\beta:0}\ {}_{h+\beta}\eta_0^{j-1} + \sum_{\beta=0}^{j-h-1} \sum_{\nu=0}^{\rho} \sum_{w=1}^{p} {}_hJ^{\rho,\tau,j-1}_{\beta:\nu,w}\ {}_{h+\beta}\eta^{j-1}_{\nu p+w},$$

where

$$
{}_hJ^{\rho,\tau}_{\beta:0} = \sum_{\alpha=\beta}^{\rho} C_\beta^{h+\beta}(-1)^{\alpha+\beta} l_{\alpha,\beta}\,\delta^{\alpha}_{(\rho-\alpha)p+\tau}, \tag{12a}
$$

$$
{}_hJ^{\rho,\tau,j-1}_{\beta:\nu,w} = \sum_{\alpha=\beta}^{\rho-\nu}\sum_{\lambda=\nu}^{\rho-\alpha} C_\beta^{h+\beta} C_{\lambda-\nu}^{-(\nu p+w)/p}(-1)^{\alpha+\beta} l_{\alpha,\beta}\,\delta^{\lambda-\nu+\alpha}_{(\rho-\alpha-\lambda)p+\tau-w}. \tag{12b}
$$

The following definition will be introduced.

DEFINITION. *A number* δ_i^ν (*cf.* (10)) *is of index* σ *provided* $i=(\sigma-\nu)p+\zeta$, *where* $0\leqq\zeta\leqq p-1$ *and* $\sigma\geqq\nu$.

In view of the assumed existence of a set of ϕ solutions (1), for which (1b) holds, the characteristic equation (5; §8) will have $\rho=0$ as a root whose multiplicity is precisely ϕ. As a consequence,

$$
\sum_k d_{n-k,0}k^i = \delta_0^i = 0 \qquad (i = 0, 1, \cdots, \phi-1), \tag{13}
$$

but

$$
\sum_k d_{n-k,0}\,k^\phi = \delta_0^\phi \neq 0. \tag{13a}
$$

Thus, not all the δ_i^ν of index ϕ are zero. It will now be proved that *the* δ_i^ν *with indices* $0, 1, \cdots, \phi-1$ *are necessarily all zero.*

We have

$$
{}_{j-1}F^{j-1}_{0,0} = {}_{j-1}J^{0,0}_{0:0}\ {}_{j-1}\eta_0^{j-1} = 0 \qquad (j = 1, \cdots, \phi),
$$

where ${}_{j-1}J^{0,0}_{0:0}=\delta_0^0$. As is seen from (13), $\delta_0^0=0$; but, apart from that, δ_0^0 would necessarily have to be zero since the numbers (1c) are not all zero. It follows that

$$
{}_hJ^{0,0}_{0:0} = 0 \qquad (h = j-1, \cdots, 0). \tag{14}
$$

Suppose

$$
{}_hJ^{0,i}_{0,0} = 0 \tag{14a}
$$

$$
(h = j-1, \cdots, 0;\ i = 0, 1, \cdots, \tau-1;\ 1 \leqq \tau \leqq p-1).
$$

In (14) the truth of (14a) has been established for $\tau=1$. On making use of (14a) we obtain

$$
{}_{j-1}F^{j-1}_{0,\tau} = {}_{j-1}J^{0,\tau}_{0:0}\ {}_{j-1}\eta_0^{j-1} = 0.
$$

Thus, in view of (12a),

(14b) $$ {}_{h}J^{0,\tau}_{0:0} = 0 \qquad (h = j-1, \cdots, 0).$$

It follows by induction that (14b) holds for $\tau = 0, \cdots, p-1$. Assume now, more generally, that

(15) $$ {}_{h}J^{i,\tau}_{\beta:0} = 0 $$
$$(h = j-1, \cdots, 0;\ \beta = 0, 1, \cdots, i;\ \tau = 0, 1, \cdots, p-1),$$

for $i = 0, \cdots, \rho-1$ $(1 \leqq \rho \leqq \phi - 1)$. For $\rho = 1$ (15) has been proved in (14b). On taking equations (15; $h=0$) in succession for $\beta = i$, $\beta = i-1, \cdots, 0$, and at each step using the relations previously obtained, we find that

(15a) $$ l_{i,i}\overset{i}{\delta}_{\tau} = l_{i-1,i-1}\overset{i-1}{\delta}_{p+\tau} = \cdots = l_{0,0}\overset{0}{\delta}_{ip+\tau} = 0 $$
$$(\tau = 0, 1, \cdots, p;\ i = 0, 1, \cdots, \rho-1).$$

Accordingly, (15) *implies that all the* δ_i^{ν} *whose indices are* $0, 1, \cdots, \rho-1$ *are zero.*

In the sequel use will be made of the fact that the δ_i^{ν} in (12a) are of index ρ and that in (12b) the δ_i^{ν} are of index $\rho-\nu-1$, for $w > \tau$, and the δ_i^{ν} are of index $\rho-\nu$, when $w \leqq \tau$. In deriving further consequences of (15) consider

(16) $$ {}_{j-1}F^{j-1}_{\rho,\tau} = 0, \quad {}_{j-2}F^{j-1}_{\rho,\tau} = 0, \cdots, \quad {}_{j-\rho-1}F^{j-1}_{\rho,\tau} = 0 \qquad (\tau = 0, 1, \cdots, p-1).$$

Whenever necessary j will be taken $\geqq \rho - 1$. By (15)

$$ {}_{j-1}F^{j-1}_{\rho,0} = {}_{j-1}J^{\rho,0}_{0:0}\ {}_{j-1}\eta_0^{j-1} $$

so that necessarily

(16a) $$ {}_{h}J^{\rho,0}_{0:0} = 0 \qquad (h = j-1, \cdots, 0).$$

On assuming, more generally,

(16b) $$ {}_{h}J^{\rho,0}_{\beta:0} = 0 $$
$$(h = j-1, \cdots, 0;\ \beta = 0, 1, \cdots, H-1;\ 1 \leqq H \leqq \rho),$$

it is observed that

$$ {}_{j-1-H}F^{j-1}_{\rho,0} = {}_{j-1-H}J^{\rho,0}_{H:0}\ {}_{j-1}\eta_0^{j-1} + \sum_{\beta=0}^{H-1} {}_{j-1-H}J^{\rho,0}_{\beta:0}\ {}_{j-1-H+\beta}\eta_0^{j-1} $$
$$ + \sum_{\beta=0}^{H} \sum_{\nu=0}^{\rho-1} \sum_{w=1}^{p} {}_{j-1-H}J^{\rho,0,j-1}_{\beta:\nu,w}. $$

In consequence of (16b) and since the δ_i^{ν} in the ${}_{h}J^{\rho,0,j-1}_{\beta:\nu,w}$ have indices $\leqq \rho - 1$,

it follows, in view of the preceding italicized statement, that

$$ {}_{j-1-H}F^{j-1}_{\rho,0} = {}_{j-1-H}J^{\rho,0}_{H:0}\;{}_{j-1}\eta_0^{j-1} = 0. \tag{16c}$$

(16c) implies that

$$ {}_{h}J^{\rho,0}_{H:0} = 0 \qquad (h = j-1, \cdots, 0). \tag{16d}$$

As (16d) follows from (16b),

$$ {}_{h}J^{\rho,0}_{\beta:0} = 0 \qquad (h = j-1, \cdots, 0;\ \beta = 0, 1, \cdots, \rho). \tag{17}$$

Assume that, for $\tau = 0, 1, \cdots, \sigma-1$ $(1 \leqq \sigma \leqq p-1)$,

$$ {}_{h}J^{\rho,\tau}_{\beta:0} = 0 \qquad (h = j-1, \cdots, 0;\ \beta = 0, 1, \cdots, \rho); \tag{18}$$

in (17) these relations have been proved for $\sigma = 1$. In view of (12a), equations (18), when examined in succession for $\beta = \rho, \rho-1, \cdots, 0$ (any h), will be seen to imply

$$ \delta^{\rho}_{\tau} = \delta^{\rho-1}_{p+\tau} = \cdots = \delta^{0}_{\rho p+\tau} = 0 \qquad (\tau = 0, 1, \cdots, \sigma - 1). \tag{18a}$$

A further consequence would be the relations

$$ {}_{j-1}J^{\rho,\sigma,j-1}_{0:0,w} = 0 \qquad (1 \leqq w \leqq \sigma). \tag{18b}$$

On making use of (15a) and (18b) it is observed that

$$ {}_{j-1}F^{j-1}_{\rho,\sigma} = {}_{j-1}J^{\rho,\sigma}_{0:0}\;{}_{j-1}\eta_0^{j-1} = 0$$

so that

$$ {}_{h}J^{\rho,\sigma}_{0:0} = 0 \qquad (h = j-1, \cdots, 0). \tag{19}$$

Assume now a set of relations, more general than (19),

$$ {}_{h}J^{\rho,\sigma}_{\beta:0} = 0 \qquad (h = j-1, \cdots, 0), \tag{19a}$$

where $\beta = 0, 1, \cdots, H-1$ $(1 \leqq H \leqq p)$. On account of (18a) and (19a) and by virtue of certain previously established facts, the relations

$$ {}_{j-1-H}F^{j-1}_{\rho,\sigma} = {}_{j-1-H}J^{\rho,\sigma}_{H:0}\;{}_{j-1}\eta_0^{j-1} = 0$$

would then follow; that is,

$$ {}_{h}J^{\rho,\sigma}_{H:0} = 0 \qquad (h = j-1, \cdots, 0).$$

Hence, by induction,

$$ {}_hJ^{\rho,\sigma}_{\beta:0} = 0 \qquad (h = j-1, \cdots, 0) \tag{20}$$

for $\beta=0, 1, \cdots, \rho$. On noting that (20) is a consequence of (18), it is observed that an induction in a more extended sense has been completed. Thus, for $\tau=0, 1, \cdots, p$,

$$ {}_hJ^{\rho,\tau}_{\beta:0} = 0 \qquad (h = j-1, \cdots, 0; \beta = 0, 1, \cdots, \rho). \tag{21}$$

Therefore, it is observed that if (15) holds for $i=0, \cdots, \rho-1$, as originally assumed, necessarily (15) will also hold for $i=\rho$. Whence it follows that

$$ {}_hJ^{i,\tau}_{\beta:0} = 0 \tag{22}$$

$$(h = j-1, \cdots, 0; \beta = 0, 1, \cdots, i; \tau = 0, 1, \cdots, p-1)$$

for $i=0, 1, \cdots, \phi-1$. Just as (15a) has been established on the basis of (15), we now conclude that all the δ_i^ν, whose indices are $0, 1, \cdots, \phi-1$, are zero. Thus, the italicized statement following (13a) has been proved.

In view of the fact just established, we obtain from (12a) and (12b)

$$ {}_hJ^{\rho,\tau}_{\beta:0} = 0 \qquad (\rho \leqq \phi - 1), \tag{23}$$

$$ {}_hJ^{\rho,\tau,j-1}_{\beta:\nu,w} = 0 \begin{cases} \rho \leqq \phi - 1 & (w \leqq \tau), \\ \rho \leqq \phi & (w > \tau). \end{cases} \tag{23a}$$

Thus, relations (8) are all satisfied for $\rho=0, 1, \cdots, \phi-1$, without yielding any information about the coefficients ${}_h\eta_s^{j-1}$.

In view of the statement preceding (16),

$$ {}_hJ^{\rho,\tau,j-1}_{\beta:\nu,w} = 0 \begin{cases} \nu \geqq \rho - \phi + 1 & (w \leqq \tau), \\ \nu \geqq \rho - \phi & (w > \tau). \end{cases} \tag{24}$$

Accordingly equations (8), which need to be considered only for $\rho\geqq\phi$, can be written in the form

$$\begin{aligned} K_h^{\rho,\tau,j-1}\ {}_h\eta^{j-1}_{(\rho-\phi)p+\tau} \equiv & \sum_{\beta=0}^{j-h-1} {}_hJ^{\rho,\tau}_{\beta:0}\ {}_{h+\beta}\eta_0^{j-1} + \sum_{\beta=0}^{j-h-1} \sum_{\nu=0}^{\rho-\phi-1} \sum_{w=1}^{p} {}_hJ^{\rho,\tau,j-1}_{\beta:\nu,w}\ {}_{h+\beta}\eta^{j-1}_{\nu p+w} \\ & + \sum_{\beta=0}^{j-h-1} \sum_{w=1}^{\tau} {}_hJ^{\rho,\tau,j-1}_{\beta:\rho-\phi,w}\ {}_{h+\beta}\eta^{j-1}_{(\rho-\phi)p+w} \end{aligned} \tag{25}$$

$$(\rho = \phi, \phi+1, \cdots; h = j-1, \cdots, 0; \tau = 0, 1, \cdots, p-1),$$

where

$$ K_h^{\rho,\tau,j-1} = -\, {}_hJ^{\rho,\tau,j-1}_{0:\nu-\phi,\tau} = - \sum_{\lambda=\rho-\phi}^{\rho} C_{\lambda-\rho+\phi}^{-[(\rho-\phi)p+\tau]/p}\, \delta^{\lambda-\rho+\phi}_{(\rho-\lambda)p}. \tag{25a}$$

In the second member of (25) the term containing ${}_h\eta^{j-1}_{(\rho-\phi)p+\tau}$ is omitted. From (25), with $\rho=\phi$, $h=j-1$, $\tau=0$, it follows that necessarily

$$\overset{0}{\delta}_{\phi p} = 0. \tag{26}$$

The essential fact is noted that in view of (13a) not all the terms in the expression for $K_h{}^{\rho,\tau,j-1}$ may be zero. It is clear, then, that the *coefficient of* η^{j-1} *in the left member of* (25) *may vanish only for a finite number of values of the involved subscripts and superscripts*. Equations (25) may be solved according to the scheme

$$\rho = \phi \left\{ \begin{array}{lcr} h = j-1\,(\tau = 0, & \cdots & , p-1) \\ h = j-2\,(\cdots & & \cdots) \\ \cdots & \cdots & \cdots \\ h = 0 \quad (\cdots & & \cdots) \end{array} \right| ; \rho = \phi + 1 \left\{ \quad \cdots \quad \right| ; \rho = \phi + 2 \cdots. \tag{27}$$

In connection with the possible vanishing of some of the numbers defined by (25a), a statement can be made precisely analogous to that preceding Lemma 1 (§3).

Lemma 6. *Consider a root, of multiplicity* ϕ, *of the characteristic equation* (5, §8) *associated with the difference equation* (B_1). *In order that, corresponding to this root, there should exist a linearly independent set of* ϕ *formal solutions of type* (1), (1a), (1b) *the following conditions are necessary and sufficient.*

(i) $$\overset{\phi}{\delta}_0 \neq 0, \qquad \overset{0}{\delta}_{\phi p} = 0.$$

(ii) *All the* $\delta_i^{\,r}$, *whose indices* (*cf. Definition*) *are* $0, 1, \cdots, \phi-1$, *are zero.*

(iii) *If any of the* $K_h{}^{\rho,\tau,j-1}(\rho \geqq \phi)$, *defined by* (25a), *are zero* (*cf. italics preceding* (27)) *then the* $\delta_i^{\,r}$ *satisfy conditions implied by the vanishing of the corresponding second members of* (25).

10. The mixed system of difference equations. The formal solutions (1; §9), (1a; §9), (1b; §9) may be written in the form

$$y_j(x)(= s_j(x)) = \sum_{h=0}^{j-1} \log^h x \left[{}_h\eta_0^{j-1} + \sum_{w=1}^{p} x^{(p-w)/p}\, {}_h\eta_w^{j-1}(x) \right] \tag{1}$$

$$(j = 1, \cdots, \phi),$$

where

$${}_h\eta_w^{j-1}(x) = \sum_{\lambda=0}^{\infty} {}_h\eta_{\lambda p+w}^{j-1} x^{-\lambda-1}. \tag{1a}$$

The coefficients in the series (1a) are specified as in §9. Let j for the present have a fixed value. Consider (1) as a transformation to be applied on the difference equation (B_1). The coefficients of (B_1) may be written in the form

$$d_{n-k}(x) = d_{n-k,0} + \sum_{\nu=1}^{p} x^{(p-\nu)/p} d_{n-k,\nu}(x), \tag{2}$$

$$d_{n-k,\nu}(x) = \sum_{\lambda=0}^{\infty} d_{n-k,\lambda p+\nu} x^{-\lambda-1}. \tag{2b}$$

From (1), in view of (3b; §9), it follows that

$$y_j(x+k) = Y^j_{k:1} + Y^j_{k:2}, \tag{3}$$

where

$$Y^j_{k:1} = \sum_{h=0}^{j-1} \log^h x \sum_{\beta=0}^{j-1-h} \sum_{\alpha=0}^{\infty} C_\beta^{h+\beta} (-1)^{\alpha+\beta} l_{\alpha,\beta} k^\alpha \; {}_{h+\beta}\eta_0^{j-1} x^{-\alpha}, \tag{3a}$$

$$Y^j_{k:2} = \sum_{h=0}^{j-1} \sum_{w=1}^{p} \log^h x \; x^{(p-w)/p} \sum_{\beta=0}^{j-1-h} {}_hA^{j-1}_{k:\beta,w} \; {}_{h+\beta}\eta_w^{j-1}(x). \tag{3b}$$

In (3b)

$${}_hA^{j-1}_{k:\beta,w} = \sum_{\alpha=0}^{\infty} \sum_{\delta=0}^{\infty} C_\beta^{h+\beta} C_\delta^{(p-w)/p} (-1)^{\alpha+\beta} l_{\alpha,\beta} k^{\alpha+\delta} x^{-\alpha-\delta}. \tag{3c}$$

On making use of (7; §4) it is observed that, by virtue of (2), (2b), (3), (3a), (3b), (3c),

$$L(y_j(x)) \equiv L_1^j + L_2^j. \tag{4}$$

Here

$$L_1^j = \sum_{h=0}^{j-1} \sum_{w=1}^{p} x^{(p-w)/p} \log^h x \sum_{\beta=0}^{j-1-h} \sum_{k=0}^{n} {}_{h,w}L^{j-1}_{\beta,k}, \tag{4a}$$

where

$$\begin{aligned} {}_{h,w}L^{j-1}_{\beta,k} &= \sum_{\zeta=1}^{w-1} \sum_{\alpha,\delta,\lambda=0}^{\infty} C_\beta^{h+\beta} C_\delta^{(p-\zeta)/p} (-1)^{\alpha+\beta} l_{\alpha,\beta} k^{\alpha+\delta} d_{n-k,\lambda p+w-\zeta} \\ &\quad \times x^{-\alpha-\delta-\lambda} \; {}_{h+\beta}\eta_\zeta^{j-1}(x+k) + \sum_{\zeta=w}^{p} \sum_{\alpha,\delta,\lambda=0}^{\infty} C_\beta^{h+\beta} C_\delta^{(p-\zeta)/p} (-1)^{\alpha+\beta} l_{\alpha,\beta} k^{\alpha+\delta} \\ &\quad \times d_{n-k,\lambda p+p+w-\zeta} x^{-\alpha-\delta-\lambda-1} \; {}_{h+\beta}\eta_\zeta^{j-1}(x+k) + \sum_{\alpha,\delta=0}^{\infty} C_\beta^{h+\beta} C_\delta^{(p-w)/p} \\ &\quad \times (-1)^{\alpha+\beta} l_{\alpha,\beta} k^{\alpha+\delta} d_{n-k,0} x^{-\alpha-\delta} \; {}_{h+\beta}\eta_w^{j-1}(x+k); \end{aligned} \tag{4b}$$

and

$$
\begin{aligned}
L_2^j = {} & \sum_{h=0}^{j-1}\sum_{w=1}^{p} \log^h x \; x^{(p-w)/p} \sum_{\alpha,\lambda=0}^{\infty}\sum_{\beta=0}^{j-1-h} C_\beta^{h+\beta}\; {}_{h+\beta}\eta_0^{j-1}(-1)^{\alpha+\beta} l_{\alpha,\beta} \\
& \times \delta^{\alpha}_{\lambda p+w} x^{-\alpha-\lambda-1} + \sum_{h=0}^{j-1} \log^h x \sum_{\alpha=0}^{\infty}\sum_{\beta=0}^{j-1-h} C_\beta^{h+\beta}\; {}_{h+\beta}\eta_0^{j-1}(-1)^{\alpha+\beta} l_{\alpha,\beta}\delta_0^{\alpha} x^{-\alpha}.
\end{aligned} \tag{4c}
$$

On writing

$$L(y_j(x)) \equiv \sum_{h=0}^{j-1}\sum_{w=1}^{p} \log^h x \;\; x^{(p-w)/p}\;\; {}_jW_{h,w},$$

it is observed that, in view of the involved formal facts, the relations ${}_jW_{h,w}=0$ ($h=0,\cdots,j-1$; $w=1,\cdots,p$) have to be satisfied; these equations may be written in the form

$$T_{h,w}^{j-1} \equiv \sum_{\beta=0}^{j-1-h}\sum_{k=0}^{n}\sum_{\zeta=1}^{p} {}_\beta a^{h,w}_{\zeta,k}(x)\;\; {}_{h+\beta}\eta_\zeta^{j-1}(x+k) = g^{h,w,j}(x) \tag{B$_2$}$$

$$(h = 0,\cdots, j-1;\; w = 1,\cdots, p).$$

Here, by virtue of (4), (4a), (4b) and (4c), the coefficients are series of the form

$${}_\beta a^{h,w}_{\zeta,k}(x) = \sum_{s=0}^{\infty} {}_\beta a^{h,w}_{\zeta,k:s} x^{-s}, \tag{5}$$

$$g^{h,w,j}(x) = -\sum_{s=1}^{\infty} g_s^{h,w,j} x^{-s}, \tag{5a}$$

where

$${}_0 a^{h,w}_{\zeta,k:0} = d_{n-k,w-\zeta} \qquad (\zeta \leqq w), \tag{6}$$

$${}_\beta a^{h,w}_{\zeta,k:0} = 0 \qquad (\zeta > w), \qquad\qquad {}_\beta a^{h,w}_{\zeta,k:0} = 0 \qquad (\beta \geqq 1;\ \zeta \leqq w), \tag{6a}$$

and, for $s \geqq 1$,

$${}_\beta a^{h,w}_{\zeta,k:s} = \sum_{\lambda=0}^{s}\sum_{\alpha=0}^{s-\lambda} C_\beta^{h+\beta} C_{s-\alpha-\lambda}^{(p-\zeta)/p}(-1)^{\alpha+\beta} l_{\alpha,\beta} d_{n-k,\lambda p+w-\zeta} k^{s-\lambda} \qquad (\zeta < w), \tag{7}$$

$$
\begin{aligned}
{}_\beta a^{h,w}_{w,k:s} = {} & \sum_{\alpha=0}^{s} C_\beta^{h+\beta} C_{s-\alpha}^{(p-w)/p}(-1)^{\alpha+\beta} l_{\alpha,\beta} k^s d_{n-k,0} \\
& + \sum_{\lambda=0}^{s-1}\sum_{\alpha=0}^{s-1-\lambda} C_\beta^{h+\beta} C_{s-1-\alpha-\lambda}^{(p-w)/p}(-1)^{\alpha+\beta} l_{\alpha,\beta} k^{s-1-\lambda} d_{n-k,(\lambda+1)p},
\end{aligned} \tag{7a}
$$

$${}_\beta a^{h,w}_{\zeta,k:s} = \sum_{\lambda=0}^{s-1}\sum_{\alpha=0}^{s-\lambda-1} C_\beta^{h+\beta} C_{s-\alpha-\lambda-1}^{(p-\zeta)/p}(-1)^{\alpha+\beta} l_{\alpha,\beta} k^{s-\lambda-1} d_{n-k,\lambda p+p-\zeta+w} \qquad (\zeta > w). \tag{7b}$$

The coefficients in (5a) may be computed with the aid of (4c). We have, on using the notation (10; §9),

$$
\begin{aligned}
g^{h,w,j}(x) &= \sum_{\alpha,\lambda=0}^{\infty} \sum_{\beta=0}^{j-1-h} C_\beta^{h+\beta} (-1)^{\alpha+\beta+1} l_{\alpha,\beta}\ {}_{h+\beta}\eta_0^{j-1} \delta_{\lambda p+w}^{\alpha} x^{-\alpha-\lambda-1} \\
&\quad + \xi(w) \sum_{\beta=0}^{j-1-h} \sum_{\alpha=0}^{\infty} C_\beta^{h+\beta} (-1)^{\alpha+\beta+1} l_{\alpha,\beta}\ {}_{h+\beta}\eta_0^{j-1} \delta_0^{\alpha} x^{-\alpha};
\end{aligned} \tag{8}
$$

here $\xi(p)=1$ and $\xi(w)=0$ when $w \neq p$. Thus,

$$
\begin{aligned}
-g_s^{h,w,j} &= \sum_{\beta=0}^{j-1-h} \sum_{\alpha=0}^{s-1} C_\beta^{h+\beta} (-1)^{\alpha+\beta+1} l_{\alpha,\beta}\ {}_{h+\beta}\eta_0^{j-1} \delta_{(s-1-\alpha)p+w}^{\alpha} \\
&\quad + \xi(w) \sum_{\beta=0}^{j-1-h} C_\beta^{h+\beta} (-1)^{s+\beta+1} l_{s,\beta}\ {}_{h+\beta}\eta_0^{j-1} \delta_0^{s}.
\end{aligned} \tag{8a}
$$

The series (5), (5a) all converge in the vicinity of infinity. Furthermore, it is seen that the "mixed" difference system (B_2) is formally satisfied by the possibly divergent series ${}_{h+\beta}\eta_t^{j-1}(x)$ (cf. (1a)).

LEMMA 7. *Write the ϕ formal solutions* ((1; §9), (1a; §9), (1b; §9)) [*corresponding to a root of multiplicity ϕ of the characteristic equation* (5; §8) *associated with the difference equation* (B_1)] *in the form* (1), (1a). *The formal series* (1a) (*with j fixed*) *will satisfy a "mixed" difference system* (B_2) *whose coefficients are given by series* (5), (5a), (6), (6a), (7), (7a), (7b), (8a), *all convergent in a neighborhood of $x=\infty$.*

11. The corresponding system of integral equations. It is clear that whenever the characteristic equation $E_1(t)=0$ (5; §8) possesses a root $t=\rho$, the numbers $\rho \pm 2\pi i\nu$ ($\nu=1, 2, \cdots$) will also be roots. In particular, this equation has roots, on the axis of imaginaries, associated with the root $t=0$. Regions P and S will be defined as follows.

Let P denote the half t-plane $Rt \geqq 0$, excluding small sectors each with vertex at $t=0$ and containing the non-zero roots of $E_1(t)=0$ in their interiors.

Let S denote a strip

$$\epsilon \leqq t_1 \leqq a;\ t_2 \geqq 0 \qquad (0 < \epsilon < a;\ t = t_1 + it_2),$$

not containing any of the roots of $E_1(t)=0$.

In the sequel, whenever the integrals

$$\int_0^\infty, \qquad \int_0^t$$

are said to be extended in P, it will be understood that the path of integration

is along a ray $\angle t = \bar{t}$, extending from the origin and situated in P. On the other hand, when these integrals are said to be extended in S, the supposition will be that the path is from $t=0$ along the positive axis of reals up to some point $t=t_1$ ($\epsilon \leqq t_1 \leqq a$); from the latter point the path will be assumed to extend, in S, along the ray $Rt = t_1$.

The variable x will be so restricted that

$$\lim_{t\to\infty} | e^{tx} t^{\alpha} | = 0 \qquad \text{(every } \alpha > 0) \tag{1}$$

along the ray $\angle t = \bar{t}$, under consideration, of P; or along a line, $Rt = t_1$, extending in S. We then have formally

$$a(x) = \sum_{s=1}^{\infty} a_s x^{-s} = \int_0^{\infty} \bar{a}(t) e^{tx} dt \qquad \text{(integration in } P \text{ or } S), \tag{2}$$

where

$$\bar{a}(t) = \sum_{\nu=1}^{\infty} \bar{a}_{\nu} t^{\nu-1}; \qquad \bar{a}_{\nu} = \frac{(-1)^{\nu} a_{\nu}}{(\nu-1)!}. \tag{2a}$$

Unless stated explicitly integrations below are in P or S.

The series (1a; §10) are formally representable as follows:

$$_{h}\eta_{w}^{j-1}(x) = \int_0^{\infty} {}_{h}\bar{\eta}_{w}^{j-1}(t) e^{tx} dt, \tag{3}$$

where

$$_{h}\bar{\eta}_{w}^{j-1}(t) = \sum_{\nu=0}^{\infty} {}_{h}\bar{\eta}_{w:\nu}^{j-1} t^{\nu}; \qquad {}_{h}\bar{\eta}_{w:\nu}^{j-1} = \frac{(-1)^{\nu+1}}{\nu!} {}_{h}\eta_{\nu p+w}^{j-1}. \tag{3a}$$

Consider now (3) as a transformation on the dependent variables to be applied to the difference system (B_2) (Lemma 7). It is not practicable to establish convergence of the series (3a) by making use of the relations (25; §9). Instead use will be made of a system of integral equations.

We have formally

$$_{h+\beta}\eta_{\zeta}^{j-1}(x+k) = \int_0^{\infty} ({}_{h+\beta}\bar{\eta}_{\zeta}^{j-1}(t) e^{kt}) e^{tx} dt \tag{4}$$

and, for $\lambda = 1, 2, \cdots$,

$$x^{-\lambda} {}_{h+\beta}\eta_{\zeta}^{j-1}(x+k) = -\int_0^{\infty} \left[\int_0^{t} \frac{(\tau - t)^{\lambda-1}}{(\lambda-1)!} e^{k\tau} {}_{h+\beta}\bar{\eta}_{\zeta}^{j-1}(\tau) d\tau \right] e^{tx} dt, \tag{4a}$$

provided

$$\left[e^{tx}\int_0^t e^{kt}\,{}_h\bar{\eta}_w^{j-1}(t)dt^{(H)}\right]_0 = 0 \tag{5}$$

$(k = 0, 1, \cdots, n;\ h = 0, \cdots, j-1;\ w = 1, \cdots, p;\ H = 1, 2, \cdots)$.*

By (2) and (2a) the second members (5a; §10) of (B_2) are expressible in the form

$$g^{h,w,j}(t) = \int_0^\infty \bar{g}^{h,w,j}(t)e^{tx}dt, \tag{6}$$

$$\bar{g}^{h,w,j}(t) = \sum_{\nu=1}^{\infty} \bar{g}_\nu^{h,w,j} t^{\nu-1}, \qquad \bar{g}_\nu^{h,w,j} = \frac{(-1)^{\nu+1}}{(\nu-1)!} g_\nu^{h,w,j}; \tag{6a}$$

moreover, by (4) and (4a),

$$\begin{aligned} {}_\beta a^{h,w}_{\zeta,k}(x)\ {}_{h+\beta}\eta_\zeta^{j-1}(x+k) = \int_0^\infty e^{tx}\Big[&{}_\beta a^{h,w}_{\zeta,k:0}\, e^{kt}\ {}_{h+\beta}\bar{\eta}_\zeta^{j-1}(t) \\ &- \int_0^t \left(\sum_{\lambda=1}^{\infty} {}_\beta a^{h,w}_{\zeta,k:\lambda} \frac{(\tau-t)^{\lambda-1}}{(\lambda-1)!} e^{k\tau}\right) {}_{h+\beta}\bar{\eta}_\zeta^{j-1}(\tau)\Big] dt. \end{aligned} \tag{7}$$

On substituting (6) and (7) in (B_2) it is found that the system (B_2) is formally satisfied if

$$\begin{aligned} &\sum_{\beta=0}^{j-h-1}\sum_{\zeta=1}^{p} {}_\beta b_\zeta^{h,w}(t)\ {}_{h+\beta}\bar{\eta}_\zeta^{j-1}(t) \\ &= \sum_{\beta=0}^{j-1-h}\sum_{\zeta=1}^{p}\int_0^t {}_\beta c_\zeta^{h,w}(t,\tau)\ {}_{h+\beta}\bar{\eta}_\zeta^{j-1}(\tau)d\tau + \bar{g}^{h,w,j}(t) \end{aligned} \tag{8}$$

$(h = 0, \cdots, j-1;\ w = 1, \cdots, p;\ j = 1, \cdots, \phi)$.

In the system (8)

$${}_\beta b_\zeta^{h,w}(t) = \sum_{k=0}^{n} {}_\beta a^{h,w}_{\zeta,k:0}\, e^{tk}, \tag{8a}$$

$${}_\beta c_\zeta^{h,w}(t,\tau) = \sum_{\lambda=1}^{\infty}\sum_{k=0}^{n} {}_\beta a^{h,w}_{\zeta,k:\lambda} \frac{(\tau-t)^{\lambda-1}}{(\lambda-1)!} e^{k\tau}, \tag{8b}$$

where the ${}_\beta a^{h,w}_{\zeta,k:\lambda}$ are given by the formulas (8), (8a), (9), (9a), (9b) of §10. In view of the convergence of the series (2b; §10) and by virtue of (8a; §10) and of (6a) it follows that

* The formal steps will be finally justified.

$$(9)\qquad \left| {}_{\beta}a_{\zeta,k:\lambda}^{h,w} \right|, \qquad \left| g_{\lambda}^{h,w,j} \right| < R\rho^{\lambda} \qquad (\lambda = 0, 1, \cdots).$$

Thus the coefficients of the system (8) are entire in the involved variables. The system (8) is formally satisfied by the series (3a).

As a consequence of (6; §10) and (6a; §10) we have

$$(10)\qquad {}_{0}b_{\zeta}^{h,w}(t) = \sum_{k=0}^{n} d_{n-k,w-\zeta}e^{kt} = b^{w-\zeta}(t) \qquad (\zeta = 1, \cdots, w),$$

$$(10\text{a})\qquad {}_{\beta}b_{\zeta}^{h,w}(t) = 0 \qquad (\zeta > w), \qquad {}_{\beta}b_{\zeta}^{h,w}(t) = 0 \qquad (\beta \geqq 1; \zeta \leqq w).$$

Accordingly the first member of (8) is replaced by

$$(11)\qquad \sum_{\zeta=1}^{w} b^{w-\zeta}(t)\ {}_{h}\bar{\eta}_{\zeta}^{j-1}(t).$$

Since, by (4a; §8), not all the $d_{n-k,0}$ are zero it is noted that $b^{0}(t) \not\equiv 0$.

It will be necessary to examine (8b) in greater detail. We have

$$(12)\qquad {}_{\beta}c_{\zeta}^{h,w}(t, \tau) = \sum_{\rho=0}^{\infty}\sum_{u=0}^{\infty} {}_{\beta}c_{\zeta:\rho,u}^{h,w} t^{\rho}\tau^{u},$$

where

$$(12\text{a})\qquad {}_{\beta}c_{\zeta:\rho,u}^{h,w} = \sum_{s=\rho+1}^{\rho+1+u}\sum_{k=0}^{n} \frac{{}_{\beta}a_{\zeta,k:s}^{h,w} C_{\rho}^{s-1}(-1)^{\rho} k^{u-s+\rho+1}}{(s-1)!(u-s+\rho+1)!}.$$

Hence by (7), (7a), (7b) of §10, on using the notation (13; §9), it follows that

$$(13)\qquad \begin{aligned} {}_{\beta}c_{w:\rho,u}^{h,w} &= \sum_{s=\rho+1}^{\rho+1+u}\sum_{\alpha=0}^{s} \frac{C_{\beta}^{h+\beta} C_{s-\alpha}^{(p-w)/p} C_{\rho}^{s-1} l_{\alpha,\beta}(-1)^{\rho+\alpha+\beta}\delta_{0}^{u+\rho+1}}{(s-1)!(u-s+\rho+1)!} \\ &+ \sum_{s=\rho+1}^{\rho+1+u}\sum_{\lambda=0}^{s-1}\sum_{\alpha=0}^{s-1-\lambda} \frac{C_{\beta}^{h+\beta} C_{s-1-\alpha-\lambda}^{(p-w)/p} C_{\rho}^{s-1} l_{\alpha,\beta}(-1)^{\rho+\alpha+\beta}\delta_{(\lambda+1)p}^{u+\rho-\lambda}}{(s-1)!(u-s+\rho+1)!}, \end{aligned}$$

and, for $1 \leqq \zeta \leqq w-1$,

$$(13\text{a})\qquad {}_{\beta}c_{\zeta:\rho,u}^{h,w} = \sum_{s=\rho+1}^{\rho+1+u}\sum_{\lambda=0}^{s}\sum_{\alpha=0}^{s-\lambda} \frac{C_{\beta}^{h+\beta} C_{s-\alpha-\lambda}^{(p-\zeta)/p} C_{\rho}^{s-1} l_{\alpha,\beta}(-1)^{\alpha+\beta+\rho}}{(s-1)!(u-s+\rho+1)!}\delta_{\lambda p+w-\zeta}^{u+\rho+1-\lambda},$$

while, for $\zeta > w$,

$$(13\text{b})\qquad {}_{\beta}c_{\zeta:\rho,u}^{h,w} = \sum_{s=\rho+1}^{\rho+1+u}\sum_{\lambda=0}^{s-1}\sum_{\alpha=0}^{s-\lambda-1} \frac{C_{\beta}^{h+\beta} C_{s-\alpha-\lambda-1}^{(p-\zeta)/p} C_{\rho}^{s-1} l_{\alpha,\beta}(-1)^{\alpha+\beta+\rho}\delta_{\lambda p+p-\zeta+w}^{u+\rho-\lambda}}{(s-1)!(u-s+\rho+1)!}.$$

Now the indices (cf. Definition of §9) of the δ_i^{ν} displayed in (13), (13a) and (13b) are $u+\rho+1$, $u+\rho+1$ and $u+\rho$, respectively. This, by virtue of the fact that condition (ii) of Lemma 6 (§9) is satisfied, implies that

$$ {}_{\beta}c_{\zeta:\rho,u}^{h,w} = 0 \qquad (\rho + u \leqq \phi - 2;\ \zeta \leqq w), \tag{14}$$

$$ {}_{\beta}c_{\zeta:\rho,u}^{h,w} = 0 \qquad (\rho + u \leqq \phi - 1;\ \zeta > w). \tag{14a}$$

As to the properties at infinity of the coefficients in the second members of (8), it is noted that by (6a) and (9)

$$ \left| \bar{g}^{h,w,j}(t) \right| < R'' e^{\rho|t|} \qquad (\text{all } t); \tag{15}$$

moreover, by (8b) and (9),

$$ \left| {}_{\beta}c_{\zeta}^{h,w}(t, \tau) \right| < R'' \left| e^{n\tau} \right| e^{\rho|t-\tau|} \qquad (\text{all } t), \tag{16}$$

provided $R\tau \geqq 0$ and R'' is sufficiently great.

It is observed that the functions (10) are representable by the series

$$ b^{w-\zeta}(t) = \sum_{i=0}^{\infty} \delta_{w-\zeta}^{i} \frac{t^i}{i!} $$

where, by Lemma 6, $\delta_{w-\zeta}^{i} = 0$ ($\zeta = 1, \cdots, w$; $i = 0, \cdots, \phi - 1$). Thus

$$ b^{w-\zeta}(t) = t^{\phi} d^{w-\zeta}(t) \qquad (\zeta = 1, \cdots, w), \tag{17}$$

the $d^{w-\zeta}(t)$ being entire in t. Furthermore, since $\delta_0^{\phi} \neq 0$, the function

$$ \frac{t^{\phi}}{b^0(t)} = \frac{t^{\phi}}{E_1(t)} = d(t) \qquad [\text{cf. } (5;\ §8);\ d(0) \neq 0] \tag{17a}$$

is meromorphic and, at $t = 0$, it is analytic.† The poles of $d(t)$ are given by the totality of the non-zero roots of the characteristic equation $E_1(t) = 0$.

On noting that by (4a; §8) $d_{m,0} \neq 0$ ($0 \leqq m$) it is concluded that, for t in P and for t in S,

$$ \left| d(t) \right| < d \left| e^{-(n-m)t} t^{\phi} \right|. \tag{18}$$

On the other hand, in view of (10) and (17),

$$ \left| d^{w-\zeta}(t) \right| < d \left| e^{nt} t^{-\phi} \right| \qquad (Rt \geqq 0). \tag{18a}$$

Substituting (10) and (10a) in the left members of (8), we bring the system (8) to the form

$$ t^{\phi}\, {}_{h}\bar{\eta}_{w}^{j-1}(t) = \sum_{\beta=0}^{j-h-1} \sum_{\zeta=1}^{p} \int_0^t {}_{\beta}\overset{*}{c}{}_{\zeta}^{h,w}(t, \tau)\, {}_{h+\beta}\bar{\eta}_{\zeta}^{j-1}(\tau) d\tau + \overset{*}{g}{}^{h,w,j}(t) \tag{B$_8$}$$

$$ (h = 0, \cdots, j - 1;\ w = 1, \cdots, p), $$

---

† With the number a, used in the definition of s, sufficiently small $d(t)$ is analytic for $|t| \leqq a$.

where

$$(19) \qquad {}^{*}_{\beta}c_{\zeta}^{h,w}(t,\tau) = d(t)\, {}_{\beta}c_{\zeta}^{h,w}(t,\tau) - d(t)d^{w-\zeta}(t)\sum_{\sigma=1}^{w-1} {}^{*}_{\beta}c_{\zeta}^{h,\sigma}(t,\tau),$$

$$(19a) \qquad {}^{*}g^{h,w,j}(t) = d(t)\tilde{g}^{h,w,j}(t) - d(t)d^{w-\zeta}(t)\sum_{\sigma=1}^{w-1} {}^{*}g^{h,\sigma,j}(t).$$

The steps to be used in this connection are precisely those employed in establishing equivalence of the system (A_3; §5) to (7; §5).

In view of the established properties of the coefficients of the system (8) it follows, by successive applications of the relations (19), (19a), that

$$(20) \qquad {}^{*}_{\beta}c_{\zeta}^{h,w}(t,\tau) = \sum_{s=0}^{\infty}\sum_{q=0}^{\infty} {}^{*}_{\beta}c_{\zeta:s,q}^{h,w}\, t^{s}\tau^{q},$$

$$(20a) \qquad \overset{*h,w,j}{g}(t) = \sum_{s=0}^{\infty} \overset{*h,w,j}{g}_{:s}\, t^{s};$$

the involved series being convergent for $|t| < \rho'$.† Moreover, the coefficients of (B_3) are meromorphic functions (in t), whose t-poles are at the non-zero roots of $E_1(t) = 0$. The functions (20) are entire in τ.

As a consequence of (14), (14a) and since $d(t)$ and the $d^{w-\zeta}(t)$ are analytic at $t=0$, it follows from (19) and (19a) that

$$(21) \qquad {}^{*}_{\beta}c_{\zeta:\rho,u}^{h,w} = 0 \begin{cases} \rho + u \leqq \phi - 2 & (\zeta \leqq w), \\ \rho + u \leqq \phi - 1 & (\zeta > w). \end{cases}$$

It will be demonstrated that, for t in P and S, and for R_0 sufficiently great,

$$(22) \qquad \left| {}^{*}_{\beta}c_{\zeta}^{h,w}(t,\tau) \right| < R_0 |t|^{\phi} \left| e^{wmt} \right| \left| e^{-n(t-\tau)} \right| e^{\rho|t-\tau|},$$

$$(22a) \qquad \left| {}^{*}g^{h,w,j}(t) \right| < R_0 |t|^{\phi} \left| e^{wmt} \right| \left| e^{-nt} \right| e^{\rho|t|}$$

($\beta = 0, 1, \cdots, j-h-1$; $\zeta = 1, \cdots, p$; $h = 0, \cdots, j-1$; $w = 1, \cdots, p$; $R\tau \geqq 0$).

In fact, by (18) and (16),

$$(23) \qquad \left| {}^{*}_{\beta}c_{\zeta}^{h,1}(t,\tau) \right| = |d(t)| \left| {}_{\beta}c_{\zeta}^{h,1}(t,\tau) \right| < dR'' |t|^{\phi} \left| e^{mt} \right| \left| e^{-n(t-\tau)} \right| e^{\rho|t-\tau|},$$

and, by (18) and (15),

$$(23a) \qquad \left| {}^{*}g^{h,1,j}(t) \right| = |d(t)| \left| \tilde{\tilde{g}}^{h,1,j}(t) \right| < dR'' |t|^{\phi} \left| e^{mt} \right| \left| e^{-nt} \right| e^{\rho|t|}.$$

Thus, provided R_0 is taken $\geqq dR''$, inequalities (22) and (22a) are seen to be

† ρ' is the least absolute value of the non-zero roots of the characteristic equation $E_1(\rho) = 0$.

true for $w=1$. Suppose that, for some R_1, for t in P and for $R\tau \geqq 0$,

$$(24)\qquad \left| {}_{\beta}\overset{*}{c}{}_{\zeta}^{h,\sigma}(t,\tau) \right| < R_1 \left| t \right|^{\phi} \left| e^{\sigma m t} \right| \left| e^{-n(t-\tau)} \right| e^{\rho|t-\tau|},$$

$$(24a)\qquad \left| {}^{*}g^{h,\sigma,j}(t) \right| < R_1 \left| t \right|^{\phi} \left| e^{\sigma m t} \right| \left| e^{-nt} \right| e^{\rho|t|}$$

$$(\sigma = 1, \cdots, w-1;\ 2 \leqq w \leqq p).$$

By (24), (24a), (18), (18a), (16), (15) from (19) and (19a) we then would have

$$\left| {}_{\beta}\overset{*}{c}{}_{\zeta}^{h,w}(t,\tau) \right| < \left| t \right|^{\phi} \left| e^{wmt} \right| \left| e^{-n(t-\tau)} \right| e^{\rho|t-\tau|} R_w(t),$$

$$\left| {}^{*}g^{h,w,j}(t) \right| < \left| t \right|^{\phi} \left| e^{wmt} \right| \left| e^{-nt} \right| e^{\rho|t|} R_w(t),$$

where, for $Rt \geqq 0$,

$$R_w(t) = d(pdR_1 + R'' \left| e^{-(w-1)mt} \right|) \leqq d(pdR_1 + R'').$$

On noting that w may assume only a finite number of values it is concluded that there exists a sufficiently great R_0 so that the inequalities (22) and (22a) all hold for $w=1, 2, \cdots, p$.

Thus, the following Lemma has been proved.

LEMMA 8. *The formal series* ${}_h\bar{\eta}_w^{j-1}(t)$ (3a), *associated with the series* (1; §9) *by means of* (1a; §9), *satisfy a certain integral system* (B_3). *The coefficients of this system are defined by convergent series* (20), (20a); *they are functions meromorphic in t, whose only finite t-singularities are poles at the non-zero roots of the characteristic equation* $E_1(t)=0$ (*cf.* (5; §8)). *The* ${}_{\beta}\overset{*}{c}{}_{\zeta}^{h,w}(t,\tau)$ *are entire in* τ. *Essential properties of these coefficients at* $(t=0, \tau=0)$ *are given by* (21). *On the other hand, for t in P and for t in S* (*cf. definitions at the beginning of this section*), *the coefficients of* (B_3) *satisfy inequalities* (22) *and* (22a).

12. **The Main Theorem for difference equations.** The series ${}_h\bar{\eta}_w^{j-1}(t)$ (3a; §11), referred to in Lemma 8 (§11), are formal solutions of a system of integral equations for which all the conditions of Lemma 4 (§7) are satisfied. Accordingly, by virtue of the latter lemma, the ${}_h\bar{\eta}_w^{j-1}(t)$ *converge in some vicinity of* $t=0$, *thus representing solutions, analytic at* $t=0$, *of the system* (B_3; §11). Such a system (B_3) exists for every $j (j=1, \cdots, \phi)$.

The following lemma will be now demonstrated.

LEMMA 9. *Suppose that a system of integral equations* (B_3; §11) *is given which satisfies the conditions of Lemma 8. The elements* ${}_h\bar{\eta}_w^{j-1}(t)$ ($h=0, \cdots, j-1$; $w=1, \cdots, p$), *associated with any particular set of solutions of* (B_3) (*they are analytic at* $t=0$), *have analytic continuations in the regions P and S*

(cf. definitions at the beginning of §11),* *for which the following inequalities hold.*

Along every line $Rt=t_1$ in S, for C and q sufficiently great and independent of the position of the line, we have

$$(1)\qquad |{}_h\bar{\eta}_w^{j-1}(t)| < Ce^{qIt} \qquad (h = 0, \cdots, j-1;\ w = 1, \cdots, p).$$

When m of (4a; §8) *is zero, then, along every ray $\angle t = \bar{l}$ in P,*

$$(1a)\qquad |{}_h\bar{\eta}_w^{j-1}(t)| < Ce^{q|t|} \qquad (h = 0, \cdots, j-1;\ w = 1, \cdots, p).$$

Consider a particular set of solutions of (B_3) and choose C so that

$$(2)\qquad |{}_h\bar{\eta}_{w:0}^{j-1}| < C \qquad (h = 0, \cdots, j-1;\ w = 1, \cdots, p),$$

and so that, for every $q \geqq 0$, we have

$$(3)\qquad |{}_h\bar{\eta}_w^{j-1}(t)| < Ce^{q|It|} \qquad (h = 0, \cdots, j-1;\ w = 1, \cdots, p),$$

when t is in S and $It \leqq \rho^0$ (ρ^0 some positive number) and also when $|t| \leqq a$. Suppose now that the part of the Lemma referring to the region S is not true. There exists then a number $\rho' > \rho^0$ such that the following holds. For $It < \rho'$, along every line $Rt = t_1$, in S,

$$(4)\qquad |{}_h\bar{\eta}_w^{j-1}(t)| < Ce^{qIt} \qquad (h = 0, \cdots, j-1;\ w = 1, \cdots, p);$$

on the other hand, for some $h = h'$, $w = w'$ and for some line $Rt = t_1'$ in S,

$$(4a)\qquad |{}_{h'}\bar{\eta}_{w'}^{j-1}(t)| = Ce^{q\rho'} \qquad (t = t' = t_1' + i\rho').$$

Consider now an integral of the second member of (B_3). Write

$$(5)\qquad \int_0^{t'} = \int_0^{t_1'} + \int_{t_1'}^{t_1'+i\rho'}.$$

It is observed that, for all involved β, ζ, h, w,

$$(6)\qquad |{}^{*}_{\beta}c_{\zeta}^{h,w}(t, \tau)| < s_1 \qquad (|t|,\ |\tau| \leqq a);$$

moreover, for $Rt \leqq a$,

$$(6a)\qquad |e^{wmt}| \leqq e^{pma}.$$

By (6) and (3)

$$(7)\qquad \left| \int_0^{t_1'} \right| < as_1C,$$

* Continuations exist, of course, in regions more extensive than P and S.

and, by (22; §11), (6a) and (4)

$$\left|\int_{t_1'}^{t_1'+i\rho'}\right| < e^{q\rho'}R_0Ce^{pma}\,|\,t'\,|^{\phi}\int_0^{\rho'} e^{-(q-\rho)(\rho'-\sigma)}d\sigma. \tag{7a}$$

From (7a), in view of (6b; §7) we further have

$$\left|\int_{t_1'}^{t_1'+i\rho'}\right| < e^{q\rho'}R_0Ce^{pma}\,|\,t'\,|^{\phi}\,\frac{1}{q-\rho}, \tag{8}$$

provided $q>\rho$. Furthermore, it is noted that

$$|\,t'\,| < \bar{a}\rho' \qquad \left(\bar{a} = \frac{1}{\rho^0}\,[a^2+(\rho^0)^2]^{1/2}\right) \tag{9}$$

where $\bar{a}$ is independent of $t'=t_1'+i\rho'$, provided t' is in S and $\rho'>\rho^0$. By (6a) and (9) from (22a) it follows that, for all involved h, w, j,

$$|\,{}^*g^{h,w,j}(t')\,| < R_0e^{pma}e^{\rho\bar{a}\rho'}\,|\,t'\,|^{\phi}. \tag{10}$$

Thus, in view of (7), (8), (5) and (10) on making use of (4a) and of (B_3), we would have

$$|\,t'\,|^{\phi}Ce^{q\rho'} < C\phi p a s_1 + CR_0\phi p e^{pma}e^{q\rho'}\,|\,t'\,|^{\phi}\,\frac{1}{q-\rho} + R_0e^{pma}e^{\rho\bar{a}\rho'}\,|\,t'\,|^{\phi}.$$

Since $\rho'>\rho^0$, this inequality would imply that

$$1 < \frac{\phi p a s_1 e^{-q\rho^0}}{(\rho^0)^{\phi}} + \frac{R_0\phi p e^{pma}}{q-\rho} + \frac{R_0}{C}\,e^{pma}e^{-(q-\rho\bar{a})\rho'} = f(q). \tag{11}$$

Now, $f(q)$ approaches zero as $q\to\infty$. Hence for q sufficiently great, so that $1\geqq f(q)$, there arises a contradiction. Accordingly, the part of the Lemma relating to the strip S has been demonstrated.

It remains to examine the case when $m=0$. We again consider a particular set of solutions of (B_3) and we take C so that (2) holds. Then, for some positive r^0 and for every $q\geqq 0$,

$$|\,{}_h\bar{\eta}_w^{j-1}(t)\,| < Ce^{q|t|} \qquad (h=0,\cdots,j-1;\ w=1,\cdots,p) \tag{12}$$

when $|t|\leqq r^0$. Assume that the part of the Lemma concerning the region P does not hold. Then there will exist a number r', $r'>r^0$, such that the following will be true. For $|t|<r'$, along every ray $\bar{t}(=\angle t)$ in P,

$$|\,{}_h\bar{\eta}_w^{j-1}(t)\,| < Ce^{q|t|} \qquad (h=0,\cdots,j-1;\ w=1,\cdots,p). \tag{13}$$

On the other hand, for some $h=h'$, $w=w'$, and for some ray $\angle t=\bar{t}'$, in P,

(13a) $$\left| {}_{h'}\bar{\eta}_{w'}^{j-1}(t) \right| = Ce^{q|t|} \qquad (t = t' = r'e^{i\bar{t}'}).$$

From (B_3), by virtue of (13), (13a), (22; §11) and (22a; §11), it would follow that

(14) $$\begin{aligned} \left| t^{\phi}\, {}_{h'}\bar{\eta}_{w'}^{j-1}(t) \right| = |t|^{\phi} Ce^{q|t|} &< p\phi \int_0^{r'} R_0 |t|^{\phi} e^{\rho(|t|-|\tau|)} Ce^{q|\tau|} d|\tau| \\ &+ R_0 |t|^{\phi} e^{\rho|t|}; \end{aligned}$$

here $t=t'$ and integration is along the ray $\angle t = \bar{t}'$. In view of (6b; §7) and since $r' > r^0$, (14) implies that

(15) $$1 < \frac{R_0 p\phi}{q-\rho} + \frac{R_0}{C} e^{-(q-\rho)r^0} = g(q).$$

On noting that $g(q) \to 0$, as $q \to \infty$, again a contradiction is seen to arise, when q is taken sufficiently great so that $1 \geqq g(q)$. Thus, the Lemma has been demonstrated completely.

Consider some line $Rt = t_1$ in S. The formal series (1a; §10) give rise to analytic functions (3; §11)

(16) $${}_h\eta_w^{j-1}(x) = \int_0^{\infty} {}_h\bar{\eta}_w^{j-1}(t) e^{tx} dt \qquad (h = 0, \cdots, j-1;\ w = 1, \cdots, p),$$

where the integration is extended in S (cf. beginning of §11) and the ${}_h\bar{\eta}_w^{j-1}(t)$ are functions of a set referred to in Lemma 9. In fact, let H denote an x-half plane for which

(17) $$R(ix) = |x| \cos\left(\frac{\pi}{2} + \bar{x}\right) < -q' < 0 \qquad (\bar{x} = \angle x;\ q' > q).$$

For x in H, in consequence of (1), the integrals (16) are seen to be absolutely convergent, when the path of integration is extended in S.

Similarly, when m of (4a; §8) is zero and integration in (16) is along a ray $\angle t = \bar{t}$ in P, the integrals (16) are observed to be absolutely convergent in a half plane $H[\bar{t}]$ (cf. (9; §7)); this fact is a consequence of (1a).

It is also noted that when x is in H condition (1; §11) is necessarily satisfied. On the other hand, for x in $H[\bar{t}]$ (ray $\bar{t}$ in P), this condition will also be satisfied, by virtue of (9; §7).

When $m=0$ the condition (5; §11) is satisfied along any ray $\angle t = \bar{t}$, in P, provided x is in $H[\bar{t}]$ and q' of (9; §7) is sufficiently great. It remains to consider the case when m is not necessarily zero. The function $f(t, x)$, within the brackets of (5; §11), vanishes for $t=0$. We have to show that

$$\lim_{t\to\infty} f(t, x) = 0 \qquad (x \text{ in } H);$$

here $t\to\infty$ along any line $Rt=t_1$, in S. On writing

$$f(t, x) = f_1(t, x) + f_2(t, x),$$

$$f_1(t, x) = e^{tx}\int_0^{t_1} e^{kt}\, {}_h\bar{\eta}_w^{j-1}(t)dt^{(H)},$$

$$f_2(t, x) = e^{tx}\int_{t_1}^{t} e^{kt}\, {}_h\bar{\eta}_w^{j-1}(t)dt^{(H)},$$

we note that, in view of (17), $\lim f_1(t, x)=0$. As to the function $f_2(t, x)$, it is found without difficulty that, by virtue of (1) and (17),* its limit along the line $Rt=t_1$ is also zero. Thus, (5; §11) holds in S.

Accordingly, all those developments which originally were of a formal character are now seen to be justified.

Application of fundamental theorems of Nörlund, referred to in §7, is possible. In fact, with the integration extended in S, an integral (16) may be expressed as

$$(18) \qquad {}_h\eta_w^{j-1}(x) = {}_he_w^{j-1}(x) + {}_hf_w^{j-1}(x)e^{xt_1}$$

where

$$(18a) \qquad {}_he_w^{j-1}(x) = \int_0^{t_1} e^{tx}\, {}_h\bar{\eta}_w^{j-1}(t)dt$$

$[0<\epsilon\leqq t_1\leqq a$; ϵ a constant used in the definition of $S]$ and

$$(18b) \qquad {}_hf_w^{j-1}(x) = \int_{t_1}^{t_1+i\infty} e^{(t-t_1)x}\, {}_h\bar{\eta}_w^{j-1}(t)dt.$$

The function (18a) is entire. On the other hand, in view of Nörlund's results and in view of the properties established, in S, for the ${}_h\bar{\eta}_w{}^{j-1}(t)$, the ${}_hf_w{}^{j-1}(x)$ are seen to be expressible by series

$$(19) \qquad {}_hf_w^{j-1}(x) = \sum_{s=0}^{\infty} \frac{{}_hf_w^{j-1}(t_1)}{x(x+i\gamma)(x+2i\gamma)\cdots(x+si\gamma)}$$

$$(\gamma > 0, \text{ sufficiently great}),$$

convergent in a plane H (cf. (17)); that is, convergent for

$$(20) \qquad Ix > H > 0 \qquad (H \text{ sufficiently great}).$$

* And also because along the line $Rt=t_1$, in S, $\angle t\to\pi/2$.

When m of (4a; §8) is zero (that is, when the solutions under considerations are associated with the μ-group (cf. §8) whose μ's are greatest) consideration of integrals (16) (with integrations along a ray in P) leads to convergent factorial series developments analogous to those obtained in §7.

Thus, *we have found a set of analytic solutions* ${}_h\eta_w{}^{j-1}(x)(h=0, \cdots, j-1;$ $w=1, \cdots, p)$ *of the "mixed" difference system* (B_2) (§10). Such results are established for $j=1, \cdots, \phi$. Now the ϕ formal solutions, under consideration, of (B_1) (§8) are given by (1; §10). In view of the established "summability" of the ${}_h\eta_w{}^{j-1}(x)$, the relations (1; §10) yield analytic expressions of a set of ϕ linearly independent solutions of (B_1). Consequently, we may formulate as follows the Main Theorem for difference equations.

THEOREM II. *Let a difference equation* (B) (§1) *be given. Suppose that corresponding to a root* ρ_1, *of multiplicity* ϕ, *of one of the associated characteristic equations* (*there is one such equation for each* μ-*group*; *cf.* §8) *there exists a linearly independent set of* ϕ *formal series solutions, all of normal type* (§8) *and all forming one logarithmic group. Bring* (B) *to the corresponding form* (B_1) (§8) *and let* $E_1(\rho)=0$ (5; §8) *be the characteristic equation, just referred to, of* (B_1). *For every* t_1, *such that* $\epsilon \leqq t_1 \leqq a$ $(0<\epsilon<a$; *a sufficiently small*), *the following is true.*

(B_1) *possesses a set of* ϕ *linearly independent analytic solutions*

$$(21)\quad y_j(x) = e^{Q(x)} x^{r} \sum_{h=0}^{j-1} \log^h x \left[{}_h\eta_0^{j-1} + \sum_{w=1}^{p} x^{(p-w)/p} \left({}_he_w^{j-1}(x) + {}_hf_w^{j-1}(x) e^{xt_1} \right) \right]$$
$$(j = 1, \cdots, \phi;\ Q(x) \text{ a polynomial in } x^{1/p}),$$

where the ${}_he_w{}^{j-1}$ *are entire functions of the type* (18a), *while the* ${}_hf_w{}^{j-1}(x)$ *are factorial series of the form* (19), *convergent for* $Ix>H>0$ (*H sufficiently great*). *The functions*

$${}_he_w^{j-1}(x) + {}_hf_w^{j-1}(x) e^{xt_1}$$

are expressible by convergent Laplace integrals of the form (16). *There exist corresponding developments for a half plane* $Ix<-H_1<0$.

When the solutions, under consideration, correspond to the μ-*group whose* μ'*s are the greatest, the following will hold for every* $\bar{l}(-\pi/2<\bar{l}<\pi/2)$, *not coincident with a value of an angle of a non-zero root of* $E_1(\rho)=0$.*

(B_1) *possesses a set of* ϕ *linearly independent analytic solutions of the form* (11; §7), *where* $Q(x)$ *contains no powers of x higher than the first. Similar developments exist for* $\pi/2<\bar{l}<3\pi/2$.

* It is supposed that $\rho_1=0$; this entails no loss of generality.

In so far as the corresponding solutions of (B) are concerned, we need only to adjoin a suitable factor exp $(\mu x \log x)$ to the expressions given in the above theorem.

The theorem is not capable of extension in the sense that even normal formal solutions of (B) do not in all cases lead to convergent factorial series developments, if corresponding to the multiple root, in question, there is more than one logarithmic group. This point will be demonstrated by means of the following example.

Let $L(y)=0$ be an equation (B_1) (§8) of third order, with $p=1$ and with all the μ_i (§8) zero. We shall take

$$d_{0,0} = 1, \quad d_{1,0} = -3, \quad d_{2,0} = 3, \quad d_{3,0} = -1, \tag{22}$$

$$d_{0,1} = d_{2,1} = 0, \quad d_{1,1} = 1, \quad d_{3,1} = -1 \tag{22a}$$

and

$$d_{3-k,s} = 0 \qquad (k = 0, 1, 2, 3; s = 2, 3, \cdots), \tag{22b}$$

except that $d_{3,3}=b<0$. This equation has a single characteristic equation with a triple root $\rho=0$. By direct substitution it can be verified that $L(y)$ is satisfied by a formal normal solution*

$$\begin{aligned} y(x) &= {}_0\overset{0}{\eta}(x) = 1 + {}_0\overset{0}{\eta}_1(x), \\ {}_0\overset{0}{\eta}_1(x) &= \sum_{\lambda=0}^{\infty} {}_0\overset{0}{\eta}_{\lambda+1} x^{-(\lambda+1)}. \end{aligned} \tag{23}$$

It can be shown that the number of logarithmic groups is greater than one.

Corresponding to (B_2) (§10) we have

$$\overset{0}{T}_{0,1} \equiv \sum_{k=0}^{3} {}_0\overset{0,1}{a}_{1,k}(x)\ {}_0\overset{0}{\eta}_1(x+k) = g^{0,1,1}(x), \tag{24}$$

where

$${}_0\overset{0,1}{a}_{1,k}(x) = \sum_{s=0}^{\infty} {}_0\overset{0,1}{a}_{1,k:s}\, x^{-s}, \qquad {}_0\overset{0,1}{a}_{1,k:s} = d_{3-k,s}, \tag{24a}$$

$$g^{0,1,1}(x) = -\sum_{s=1}^{\infty} g_s^{0,1,1} x^{-s}; \tag{24b}$$

here

$$g_s^{0,1,1} = 0 \qquad (s \neq 3), \qquad g_3^{0,1,1} = b. \tag{24c}$$

---

* Use is made of a notation conforming with that employed in §§8, 9, 10, 11, 12.

Let $\angle t = \bar{t}$ be any ray, extending from $t=0$ and not coincident with either half of the axis of t-imaginaries. Let integrations be along such a ray. Corresponding to (3; §11) we write, formally,

$$(25) \qquad {}_0\eta_1^0(x) = \int_0^\infty {}_0\bar{\eta}_1^0(t)e^{tx}dt \qquad (R(e^{i\bar{t}}x) < -q' < 0),$$

where q' is sufficiently great and

$$(25a) \qquad {}_0\bar{\eta}_1^0(t) = \sum_{\nu=0}^{\infty} \eta_\nu t^\nu \qquad \left(\eta_\nu = \frac{(-1)^{\nu+1}}{\nu!}\, {}_0\eta_{\nu+1}^0\right).$$

If it were demonstrated that (25a) diverges, impossibility of representing $y(x)$ in terms of a convergent factorial series would have been established. This follows by a reasoning analogous to that employed for a similar purpose at the end of §7.

Use of the difference equation appears to be impracticable in proving divergence of the series (25). However, a certain integral equation will serve this purpose. Substitution of (25) in (24) (compare with (8; §11)) leads to the equation

$$(26) \qquad (e^t - 1)^3\, {}_0\bar{\eta}_1^0(t) = \int_0^t \left(e^{2\tau} - 1 + \frac{b}{2}(\tau - t)^2\right) {}_0\bar{\eta}_1^0(\tau)d\tau + \frac{b}{2}t^2$$

which yields the following relations for the coefficients of the series (25a):

$$(27) \qquad \frac{1}{r}\eta_{r-2} = \sum_{H=0}^{r-3} f_H(r)\eta_H \qquad (r \geqq 3), \qquad \eta_0 = -\frac{b}{2},$$

where

$$(27a) \qquad f_{r-3}(r) = \frac{5}{6} - \frac{2}{r} - \frac{b}{r(r-1)(r-2)},$$

$$f_H(r) = \frac{1}{(r-H)!}\delta_0^{r-H} - \frac{1}{(r-H-1)!r!}2^{r-H-1} \quad (0 \leqq H \leqq r-4).$$

Since $b<0$, $f_{r-3}(r)>0$ $(r \geqq 3)$. We have $\delta_0^{r-H} = 3^{r-H} - 3(2^{r-H}) + 3$. Thus, for $0 \leqq H \leqq r-4$,

$$(27b) \qquad f_H(r) = \frac{3^{r-H-1}}{(r-H)!}(f_H'(r) - f_H''(r))$$

where

$$f_H'(r) = 3\left[1 - 2\left(\frac{2}{3}\right)^{r-H-1} + \frac{1}{3^{r-H-1}}\right] > \frac{11}{9},$$

$$f_H''(r) = \frac{r-H}{r}\left(\frac{2}{3}\right)^{r-H-1} \leqq \frac{8}{27}.$$

Accordingly, from (27b) it follows that, for $0 \leqq H \leqq r-4$,

$$f_H(r) = \frac{3^{r-H-1}}{(r-H)!}\zeta_H(r), \qquad \zeta_H(r) > \frac{25}{27}.$$

With the coefficients $f_H(r)$ in the second members of (27) all positive, it is seen that the η_ν of (25a) are uniquely determined, positive numbers. Hence relations (27) would imply that

$$\frac{\eta_{r-2}}{r} > f_{r-3}(r)\eta_{r-3} \qquad (r \geqq 4). \tag{28}$$

Now, by (27a), $rf_{r-3}(r) \to \infty$ as $r \to \infty$. Consequently, inequalities (28) lead to the conclusion that the series (25a) diverges for all $t(\neq 0)$. Thus, the normal formal solution of the example under consideration cannot be represented in terms of a convergent factorial series.

UNIVERSITY OF ILLINOIS,
URBANA, ILL.

SUFFICIENT CONDITIONS IN THE PROBLEM OF LAGRANGE WITHOUT ASSUMPTIONS OF NORMALCY*

BY
MARSTON MORSE

Sufficient conditions for an extremal to give a minimum in the ordinary fixed end point problem involve the Jacobi, Weierstrass, and Clebsch conditions. It has been an outstanding problem to establish the corresponding theorem in the problem of Lagrange without assumptions of normalcy or analyticity. Carathéodory [3] reduced the assumptions as to normalcy by introducing the notion of class. More recently Hestenes [4] has employed a similar notion of order of normalcy in dealing with the Jacobi conditions. The paper of Hestenes contains a number of important results independent of the assumption of normalcy.

The present paper establishes sufficient conditions involving the Jacobi, Weierstrass, and Clebsch conditions, employing for the first time, it is believed, no condition of normalcy.

In establishing the desired theorem the writer has come upon a new and powerful method of treating Mayer fields of secondary extremals. This method has also proved the proper tool in attacking other problems not involving a minimum. The fixed end point theorem is treated first and followed by the theorem for the variable end point problem in the modified Bolza [1] form.

The importance of freeing these theorems from the assumptions of normalcy is readily seen upon recalling that the theorems now established lead by simple transformations to corresponding theorems in the Mayer, parametric, and other general forms of the problem, and include earlier theorems of the same general character as special cases.

1. The functional. One is concerned with a set of functions

$$f(x, y, p), \qquad \phi_\beta(x, y, p) \qquad (\beta = 1, \cdots, m) \tag{1.1}$$

of the variables

$$x, \qquad (y) = (y_1, \cdots, y_n), \qquad (p) = (p_1, \cdots, p_n) \qquad (m < n) \tag{1.2}$$

on an open region R of the space of the variables (x, y, p). We suppose the functions (1.1) are of class C^3 on R. Our functional is of the form

* Presented to the Society, September 6, 1934; received by the editors July 3, 1934.

$$J = \int_{a^1}^{a^2} f(x, y, y')dx,$$

subject to the conditions

$$\phi_\beta(x, y, y') = 0 \qquad (\beta = 1, \cdots, m). \tag{1.3}$$

We term an element (x, y, y') *differentially admissible* if it satisfies (1.3). An arc $y_i(x)$ is termed differentially admissible if it is of class D^1 and its elements satisfy (1.3).

We set

$$F(x, y, p, \lambda) = f + \lambda_\beta \phi_\beta \qquad (\beta = 1, \cdots, m).$$

By an extremal we mean an arc of class C^2 together with multipliers $\lambda_\beta(x)$ of class C^1 which satisfy the conditions

$$\frac{d}{dx} F_{p_i} - F_{y_i} = 0, \qquad (y') = (p) \qquad (i = 1, \cdots, n),$$

and the conditions (1.3). We suppose g is such an extremal and is of the form

$$y_i = \bar{y}_i(x), \qquad \lambda_\beta = \bar{\lambda}_\beta(x)$$

for x on an interval

$$a^1 \leqq x \leqq a^2.$$

It will frequently be convenient to suppose that g is an inner segment of a slightly longer extremal. By an *admissible* arc we mean (in §§1, 2, 3, 4) a differentially admissible arc which joins the end points of g. We shall enumerate the conditions under which g affords a minimum to J relative to neighboring admissible arcs.

It will be convenient to evaluate certain functions *along* g, that is, to set

$$[x, y, p, \lambda] = [x, \bar{y}(x), \bar{y}'(x), \bar{\lambda}(x)].$$

We shall indicate such an evaluation by adding the superscript 0 to the function involved.

We assume that

$$\overset{0}{F}_{p_i p_j} z_i z_j > 0 \qquad (i, j = 1, \cdots, n) \tag{1.4}$$

for each point x on g and set $(z) \neq (0)$ for which

$$\overset{0}{\phi}_{\beta p_i} z_i = 0 \qquad (\beta = 1, \cdots, m).$$

We term this condition the *Clebsch S-condition*.

One sets

$$E[x, y, y', \lambda, Y'] = F(x, y, Y', \lambda) - F(x, y, y', \lambda) - (Y_i' - y_i')F_{p_i}(x, y, y', \lambda).$$

We assume that

$$E(x, y, y', \lambda, Y') > 0 \tag{1.5}$$

for each set (x, y, y', λ) in a neighborhood of the sets $(x, \bar{y}, \bar{y}', \bar{\lambda})$ on g and for arbitrary sets (Y'), provided merely the sets (x, y, y') and (x, y, Y') are differentially admissible and distinct. We term this condition the *Weierstrass S-condition.*

To define the third condition we set

$$2\omega(\eta, \eta') = \overset{0}{F}_{y_iy_j}\eta_i\eta_j + 2\overset{0}{F}_{y_ip_j}\eta_i\eta_j' + \overset{0}{F}_{p_ip_j}\eta_i'\eta_j' \qquad (i, j = 1, \cdots, n),$$

$$\Phi_\beta(\eta, \eta') = \overset{0}{\phi}_{\beta p_i}\eta_i' + \overset{0}{\phi}_{\beta y_i}\eta_i \qquad (\beta = 1, \cdots, m).$$

The functional

$$I = \int_{a^1}^{a^2} 2\omega(\eta, \eta')dx,$$

subject to the conditions

$$\Phi_\beta(\eta, \eta') = 0 \qquad (\beta = 1, \cdots, m),$$

is termed the *second variation.* One sets

$$\Omega(\eta, \eta', \mu) = \omega + \mu_\beta\Phi_\beta.$$

The Euler equations corresponding to the second variation take the form

$$\frac{d}{dx}\Omega_{\eta_i'} - \Omega_{\eta_i} = 0, \qquad \Phi_\beta = 0. \tag{1.6}$$

The corresponding extremals

$$\eta_i = \eta_i(x), \qquad \mu_\beta = \mu_\beta(x) \tag{1.7}$$

are termed *secondary extremals.*

It is convenient to set

$$\Omega_{\eta_i'} = \zeta_i, \qquad \Phi_\beta = 0. \tag{1.8}$$

For each value of x the equations (1.8) serve as a transformation from the variables (η, η', μ) to the variables (η, ζ). In particular the secondary extremal (1.7) can be represented in the form

$$\eta_i = \eta_i(x), \qquad \zeta_i = \zeta_i(x).$$

If the components $\eta_i(x)$ of a secondary extremal all vanish for two distinct values of x, say a and b, but are not all identically null between a and b, the values a and b are termed *conjugate*.

We shall assume that there is no value on the interval $a^1 < x \leqq a^2$ conjugate to a^1. We term this condition the *Jacobi S-condition*.

2. Anormal secondary extremals. Secondary extremals for which all the components $\eta_i(x)$ are identically null on an interval (a, b) will be termed anormal on (a, b). Other secondary extremals will be termed normal on (a, b). In particular the solution $\eta_i \equiv \zeta_i \equiv 0$ is anormal.

Let α be a number on the interval $a^1 < \alpha \leqq a^2$. Let $N(\alpha)$ be a set of secondary extremals for which $\eta_i(a^1) = 0$ and which contains the maximum number of such secondary extremals independent of secondary extremals which are anormal on (a^1, α). All secondary extremals for which $\eta_i(a^1) = 0$ will be linearly dependent on extremals anormal on (a^1, α) and extremals of $N(\alpha)$.

Let $\pi(\alpha)$ be the number of extremals in $N(\alpha)$. We observe that $\pi(\alpha)$ is monotonically increasing. There will accordingly exist at most a finite set of values of α, say $\alpha_1, \cdots, \alpha_r$, such that

$$a^1 < \alpha_1 < \cdots < \alpha_r < a^2 \tag{2.1}$$

at which $\pi(\alpha)$ is discontinuous. The integer r may in particular be null. We set

$$a^1 = \alpha_0, \qquad a^2 = \alpha_{r+1}$$

and

$$N(\alpha_h) = N_h, \qquad \pi(\alpha_h) = \pi_h \qquad (h = 1, \cdots, r + 1).$$

There will exist a set M_h of $n - \pi_h$ secondary extremals which are anormal on (a^1, α_h) which with the extremals of N_h form a set A_h of n independent secondary extremals on which $(\eta) = (0)$ at a^1. On the kth extremal of the set A_h suppose that

$$\zeta_i(a^1) = b_{ik} \qquad (k = 1, \cdots, n). \tag{2.2}$$

We suppose that the first $n - \pi_h$ extremals of the set A_h form the set M_h. Without loss of generality we can also suppose that the columns of the matrix $\|b_{ik}\|$ have been normed and orthogonalized. We introduce a set M_h' of $n - \pi_h$ secondary extremals of which the kth satisfies the conditions (Hestenes [4], §5)

$$\eta_i(a^1) = b_{ik}, \qquad \zeta_i(a^1) = 0 \qquad (k = 1, \cdots, n - \pi_h).$$

A set of n secondary extremals which are independent and mutually con-

jugate in the sense of von Escherich is called a *conjugate base*. The extremals of N_h and M_h' together form a conjugate base S_h. Let the extremals of S_h be represented by the columns of the matrix

$$\left\| \begin{matrix} \overset{h}{\eta}_{ij}(x) \\ \overset{h}{\zeta}_{ij}(x) \end{matrix} \right\| \qquad (i, j = 1, \cdots, n). \tag{2.3}$$

The determinant

$$D_h(x) = | \overset{h}{\eta}_{ij}(x) | \tag{2.3$'$}$$

is called the determinant of the conjugate base. In (2.3) it will be convenient to suppose that the columns which represent extremals of M_h' come first.

We shall prove the following lemma.

LEMMA 2.1. *The determinant of the conjugate base S_h vanishes at no point on the interval*

$$\alpha_{h-1} < x \leqq \alpha_h. \tag{2.4}$$

Let

$$\eta_i(x) \equiv c_j \overset{h}{\eta}_{ij}(x) \qquad (i, j = 1, \cdots, n) \tag{2.5}$$

be an arbitrary linear combination of the columns of the determinant (2.3)$'$. We suppose that $\eta_i(x)$ vanishes at some point x_0 on the interval (2.4), and shall prove that the constants c_j are then all null.

Let $\bar{\zeta}_i(x)$ represent the components ζ_i of the kth anormal extremal of the set M_h. One has the integral

$$\eta_i \bar{\zeta}_i \equiv \text{const}.$$

Upon making use of (2.5) and of the fact that the constants (2.2) are normed and orthogonalized we find that (Hestenes [4], §5),

$$\eta_i(a^1)\bar{\zeta}_i(a^1) \equiv c_k \equiv \eta_i(x)\bar{\zeta}_i(x).$$

We infer that $\eta_i(x)$ in (2.5) can vanish only if

$$c_k = 0 \qquad (k = 1, \cdots, n - \pi_h).$$

The columns involved in (2.5) thus belong at most to extremals of N_h, and in particular in (2.5), $\eta_i(a^1) = 0$ for each value of i. But $\eta_i(x_0) = 0$ where x_0 is on the interval (2.4). We infer that

$$\eta_i(x) \equiv 0 \qquad (a^1 \leqq x \leqq x_0) \qquad (i = 1, \cdots, n) \tag{2.6}$$

since x_0 is not a conjugate point of a^1.

We can show that

$$\eta_i(x) \equiv 0 \qquad (a^1 \leqq x \leqq \alpha_h). \tag{2.7}$$

To that end let γ represent the secondary extremal obtained by combining the extremals of N_h with the same constants as are used in (2.5). If c were the maximum value of x such that $\eta_i(x) \equiv 0$ on the interval (a^1, c), and $c < \alpha_h$, the function $\pi(\alpha)$ would be discontinuous at c. In fact for $\alpha > c$, the set $N(\alpha)$ could be taken as one which included the extremals of $N(c)$, γ, and possibly other extremals. We infer that (2.7) holds as stated.

But according to the nature of N_h the identity (2.7) is valid only if all constants c_j in (2.5) are null. The lemma follows directly.

3. Curves which are admissible relative to a conjugate family. Let K be a set of n independent mutually conjugate secondary extremals, and let L denote the set of all extremals linearly dependent on the extremals of K. Let the extremals of L be represented by giving their components η_i and multipliers μ_β as follows:

$$\begin{aligned} \eta_i &= c_j \eta_{ij}(x) \qquad (i, j = 1, \cdots, n), \\ \mu_\beta &= c_j \mu_{\beta j}(x) \qquad (\beta = 1, \cdots, m). \end{aligned} \tag{3.1}$$

We make no assumption concerning the vanishing of the determinant $|\eta_{ij}(x)|$. By the Hilbert integral *belonging* to L we mean a line integral in the space of the variables

$$(x, c) = (x, c_1, \cdots, c_n)$$

of the form

$$H = \int (\Omega - \Omega_{\eta_i'} \eta_i')dx + \Omega_{\eta_i'} d\eta_i$$

in which the variables η_i, μ_β are to be replaced by the respective right members of (3.1) and in which we set

$$\begin{aligned} \eta_i' &= c_j \eta_{ij}'(x), \\ d\eta_i &= \eta_{ij}(x)dc_j + c_j \eta_{ij}'(x)dx. \end{aligned}$$

The Hilbert integral will thus take the form

$$H = \int A(x, c)dx + B_i(x, c)dc_i. \tag{3.2}$$

The variables (c) are arbitrary and x lies on the interval (a^1, a^2). That the integral H is independent of the path in the space (x, c) follows in the usual

way from the fact that the members of the set K are mutually conjugate extremals.

The equations

$$\eta_i = c_j \eta_{ij}(x) \tag{3.3}$$

define a transformation from the space (x, c) to the space (x, η). Let

$$c_j = c_j(x) \tag{3.4}$$

represent a curve of class D^1 in the space (x, c). The image under (3.3) of a curve of the form (3.4) will be termed a curve in the space (x, η) which is *admissible relative* to L. This curve will be of class D^1, but not every curve of class D^1 in the space (x, η) will in general be the image of a curve in the space (x, c) of the form (3.4), as examples would show.

We can however prove the following theorem.

THEOREM 3.1. *If the Clebsch S-condition holds, any segment γ of a secondary extremal of a conjugate family L affords a minimum to the second variation relative to curves λ which join γ's end points and are admissible relative to L.*

In the space (x, c), γ is represented by a straight line γ_0 on which x alone varies, while λ is represented by a curve λ_0 of the form (3.4). The curve γ_0 does not necessarily join the end points of λ_0 in the space (x, c). If in particular the first end points of γ_0 and λ_0 are distinct, these end points can be joined in the space (x, c) by a straight line p on which x is constant. The line p will correspond under (3.3) to the common first end point of γ and λ. Along p the Hilbert integral H will be null. It follows that H has the same value along γ_0 as along λ_0.

Proceeding formally as in the case of ordinary Mayer fields, one sees that

$$\Delta I = I_\lambda - I_\gamma = \int_{a^1}^{a^2} E_2(x, \eta, \eta', \mu, \bar{\eta}')dx$$

where E_2 is the Weierstrass E-function for the second variation, with

$$\eta_i = c_j(x)\eta_{ij}(x), \quad \eta_i' = c_j(x)\eta_{ij}'(x) \qquad (i, j = 1, \cdots, n),$$
$$\mu_\beta = c_j(x)\mu_{\beta j}(x) \qquad (\beta = 1, \cdots, m)$$

therein, and with $\bar{\eta}_i'$ taken as the ith slope of the curve λ at the point x on λ. From the fact that the Clebsch S-condition holds it follows that E_2 is never negative for differentially admissible sets (x, η, η') and $(x, \eta, \bar{\eta}')$. We conclude that

$$\Delta I \geqq 0,$$

and the theorem is proved.

We add the following lemma.

LEMMA 3.1. *There exists a positive constant δ so small that any segment of a secondary extremal γ on which $a \leqq x \leqq a+\delta$, where a is on the interval (a^1, a^2), affords a proper minimum to the second variation relative to differentially admissible curves of class D^1 which join γ's end points.*

The proof of this lemma is readily given upon setting up a Mayer field of secondary extremals containing γ. Cf. Morse [7], Lemma 3 and Theorem 4. Such a Mayer field exists for x on the interval $(a, a+\delta)$ provided δ is a sufficiently small positive constant.

4. Fixed end points, sufficient conditions. We continue with the following lemma.

LEMMA 4.1. *In order that the second variation be non-negative for differentially admissible curves which join the end points of the segment (a^1, a^2) of the x axis, it is sufficient that the Clebsch and Jacobi S-conditions hold along g.*

We return to the notation of §2, and in particular to the constants α_k.

(e) If d is a sufficiently small positive constant, any secondary extremal γ on which x varies on an interval of the form

$$a^1 \leqq x \leqq \alpha_k + d \qquad (k = 0, 1, \cdots, r+1) \tag{4.1}$$

and on which $(\eta) = (0)$ when $x = a^1$, affords a minimum to the second variation relative to differentially admissible curves of class D^1 which join its end points.

To prove (e) we turn to the conjugate base

$$S_h \qquad (h = 1, \cdots, r+1)$$

of §2, and recall that the determinant $D_h(x)$ of this base does not vanish on the interval (2.4). There accordingly exists a positive constant d_0 independent of h such that $D_h(x)$ does not vanish on the interval

$$\alpha_{h-1} < x \leqq \alpha_h + d_0. \tag{4.2}$$

We suppose moreover that d_0 is less than the constant δ of Lemma 3.1. Statement (e) is valid if we set $d = d_0$, as we shall now prove.

Statement (e) is valid if $k=0$ by virtue of Lemma 3.1. Proceeding inductively we shall assume that (e) holds for $d = d_0$ and $k = h-1$, and shall prove that (e) holds for $d = d_0$ and $k = h$.

Let γ be a secondary extremal on which $(\eta) = (0)$ when $x = a^1$, and on which x varies on the interval (4.1) for $k = h$. Let λ be a differentially admissible curve which joins the end points of γ. Let a be the segment of λ on which

$$a^1 \leqq x \leqq \alpha_{h-1} + d_0 \tag{4.3}$$

and b the remaining segment of λ, so that we can write

$$\lambda = a + b.$$

Now any differentially admissible arc whose end points are not conjugate can be joined by an arc $\eta_i(x)$ belonging to a secondary extremal (Hestenes [4], Lemma 7.2). With this understood let τ be an arc $\eta_i(x)$ belonging to a secondary extremal and joining the end points of a. We introduce the curve

$$\mu = \tau + b.$$

The curve μ joins the end points of γ.

We are assuming that (e) holds for $d = d_0$ and $k = h-1$. Hence

$$I_a \geqq I_\tau. \tag{4.4}$$

It follows that

$$I_{a+b} \geqq I_{\tau+b}. \tag{4.5}$$

We shall now establish the inequality

$$I_{\tau+b} \geqq I_\gamma. \tag{4.6}$$

To that end we represent curves $\eta_i(x)$ belonging to secondary extremals dependent on S_h in the form

$$\eta_i = c_j \overset{h}{\eta}_{ij}(x). \tag{4.7}$$

We regard (4.7) as defining a transformation from the space (x, c) to the space (x, η). This transformation is non-singular for x on the interval (4.2). The curve b is accordingly the image in the space (x, c) of a uniquely defined curve

$$c_j = c_j(x) \qquad (j = 1, \cdots, n)$$

of class D^1 on the interval

$$\alpha_{h-1} + d_0 \leqq x \leqq \alpha_h + d_0. \tag{4.8}$$

Let c_j^0 be the value of $c_j(x)$ when $x = \alpha_{h-1} + d_0$. For x on the interval (4.3), τ will coincide with the curve of the family (4.7) determined by the constants c_j^0. The curve τ is accordingly the image under the transformation (4.7) of the straight line

$$c_j = c_j^0 \qquad (a^1 \leqq x \leqq \alpha_{h-1} + d_0)$$

in the space (x, c). The curve $\tau + b$ is thus admissible relative to the conjugate family determined by S_h. The inequality (4.6) follows from Theorem 3.1.

Combining (4.5) and (4.6) we find that

$$I_{a+b} \geqq I_\gamma. \tag{4.9}$$

Statement (e) is thereby proved.

The lemma is a consequence of statement (e) in the case where $k=r+1$.

It will follow from the next lemma that under the conditions of Lemma 4.1 the second variation is *positive definite.*

LEMMA 4.2. *If the Clebsch and Jacobi S-conditions hold along g, there exists a conjugate base of secondary extremals whose determinant does not vanish on the interval $a^1 \leqq x \leqq a^2$.*

The proof of this lemma is nearly the same as the proof of Theorem 3, Morse [7].

We start with the conjugate base S_{r+1} of §2. The determinant formed from this base does not vanish at $x=a^2$. We can accordingly choose a base

$$\left\| \begin{matrix} \eta_{ij}(x) \\ \zeta_{ij}(x) \end{matrix} \right\| \qquad (i, j = 1, \cdots, n)$$

for members of the family in which the jth column represents a member of the family such that

$$\overset{2}{\eta}_{ij} = \delta_i^j.$$

The superscripts 1 and 2 are used to indicate evaluation for $x=a^1$ and $x=a^2$ respectively. We introduce a second conjugate family H with base B of the form

$$\left\| \begin{matrix} \bar{\eta}_{ij}(x) \\ \bar{\zeta}_{ij}(x) \end{matrix} \right\|$$

and such that

$$\overset{2}{\bar{\eta}}_{ij} = \delta_i^j, \qquad \overset{2}{\bar{\zeta}}_{ij} = \overset{2}{\zeta}_{ij} - \delta_i^j.$$

We represent the family H in the form

$$\begin{aligned} \eta_i &= u_j \bar{\eta}_{ij}(x), \\ \zeta_i &= u_j \bar{\zeta}_{ij}(x), \end{aligned}$$

where the symbols u_j represent constants.

The conjugate base B will serve as the conjugate base whose existence is affirmed in the lemma.

To establish the truth of this statement we assume that it is false, and hence that the determinant $|\bar{\eta}_{ij}|$ vanishes at some point $x=c$ on the interval

$$a^1 \leqq x < a^2. \tag{4.10}$$

There will then exist an extremal γ of H determined by a set of constants $(u)\neq(0)$ such that $\eta_i(c)=0$ on γ for each i. Let λ be a curve which consists of the x axis from $x=a^1$ to $x=c$ and of the curve $\eta_i=\eta_i(x)$ belonging to γ from $x=c$ to $x=a^2$. Upon evaluating the second variation I along λ from $x=a^1$ to $x=a^2$ we find that

$$I_\lambda = \overset{2}{\zeta}_{ij} u_i u_j - u_i u_i.$$

On the other hand there will be a secondary extremal μ dependent on the base S_{r+1} for which the curve $\eta_i(x)$ joins the end points of λ. This extremal will have the form

$$\eta_i = \eta_{ij}(x)u_j, \qquad \zeta_i = \zeta_{ij}(x)u_j,$$

and we see that

$$I_\mu = \overset{2}{\zeta}_{ij} u_i u_j.$$

It follows from statement (e) of the proof of the preceding lemma that $I_\lambda \geqq I_\mu$. But this is impossible since $(u)\neq(0)$.

We infer that the determinant $|\bar{\eta}_{ij}(x)|$ does not vanish on the interval (4.10). The base B will thus serve as the base of the lemma, and the proof is complete.

We come to a basic theorem.

THEOREM 4.1. *In order that the extremal g afford a proper minimum to J relative to neighboring admissible curves, it is sufficient that the Clebsch, Jacobi, and Weierstrass S-conditions hold along g.*

The conjugate family of secondary extremals whose existence is affirmed in the last lemma forms a Mayer field of secondary extremals covering a neighborhood of the segment (a^1, a^2) of the x axis. This family can be used as in Morse [7], Theorem 4, to establish the existence of a Mayer field of primary extremals including g and covering a neighborhood of g.

The theorem follows in the usual manner.

5. General end conditions. We turn to the problem under general end conditions. The preceding results lead to a set of sufficient conditions involving the Jacobi condition which make no assumption concerning normalcy.

The form of the problem is a modification of the Bolza problem introduced by the author (Morse [8]). We suppose that we have an extremal g as before. Points near the first and last end points of g will be denoted by (x^1, y^1) and (x^2, y^2) respectively. The end conditions have the form

$$x^s = x^s(\alpha), \qquad y_i^s = y_i^s(\alpha) \qquad (s = 1, 2;\ i = 1, \cdots, n) \tag{5.1}$$

where (α) represents a set of r variables α_h. The functions $x^s(\alpha)$ and $y_i^s(\alpha)$ are assumed to be of class C^2 for (α) near (0) and to yield the end points of g when $(\alpha)=(0)$. No assumption is made concerning the rank of the matrix of the functions on the right of (5.1). The differential conditions are as before, but the functional J is replaced by the more general functional

$$J = \theta(\alpha) + \int_{x^1(\alpha)}^{x^2(\alpha)} f(x, y, y')dx$$

in which $\theta(\alpha)$ is a function of (α) of class C^2 for (α) near (0).

By an *admissible curve* γ and set (α) we mean a differentially admissible curve γ and set (α) such that γ satisfies the end conditions with the set (α). The problem is one of determining conditions under which g and the set $(\alpha)=(0)$ afford a minimum to J relative to J's value for admissible curves γ and sets (α) for which γ neighbors g and (α) neighbors (0).

We assume that g and the set $(\alpha)=(0)$ satisfy the *transversality condition* (Morse and Myers [6])

$$[(\overset{0}{F} - \overset{0}{F}_{p_i}\bar{y}_i')dx^s + \overset{0}{F}_{p_i}dy_i^s]_1^2 + d\theta \equiv 0, \tag{5.2}$$

as an identity in the differentials $d\alpha_h$ in terms of which dx^s, dy_i^s, and $d\theta$ are to be expressed.

The second variation is the functional (Morse [8])

$$I = b_{hk}\, u_h\, u_k + \int_{a^1}^{a^2} 2\omega(\eta, \eta')dx \qquad (h, k = 1, \cdots, r), \tag{5.3}$$

subject to secondary end conditions of the form

$$\overset{s}{\eta}_i = \overset{s}{c}_{ih}\, u_h \qquad (s = 1, 2;\ h = 1, \cdots, r), \tag{5.4}$$

where b_{hk} and c_{ih}^s represent constants of which $b_{hk}=b_{kh}$. It is here understood that $x^1=a^1$ and $x^2=a^2$. The differential conditions are as before. A curve γ and set (u), such that γ is differentially admissible and satisfies (5.4) with the set (u), is termed *admissible.*

For a problem under general end conditions Mayer has stated a sufficiency condition in terms of a quadratic form. Bliss [2] and Hestenes [4]

have modified this condition. We introduce a further modification of this condition which simplifies its use.

To that end let

$$(5.5)\qquad \eta_{ip}(x),\quad \zeta_{ip}(x),\quad u_{hp}\qquad (i = 1, \cdots, n;\ h = 1, \cdots, r;\ p = 1, \cdots, q)$$

be a set of q secondary extremals and constants (u) which satisfy (5.4). Suppose moreover that this set contains the maximum number of admissible secondary extremals and constants (u) which are independent of sets $\eta_i(x) \equiv 0$, $\zeta_i(x)$, $(u) = (0)$ belonging to anormal secondary extremals. Each admissible secondary extremal and set (u) is linearly dependent upon the members of the set (5.5) together with an anormal secondary extremal and set $(u) = (0)$. We consider the family

$$(5.6)\qquad \eta_i = v_p\,\eta_{ip}(x),\qquad \zeta_i = v_p\,\zeta_{ip}(x),\qquad u_h = v_p\,u_{hp},$$

of admissible secondary extremals and corresponding sets (u). Upon evaluating I along the member of this family determined by (v) one obtains a quadratic form $H(v)$. By the *Mayer S-condition* we mean the condition that $H(v)$ be positive definite.

The theorem here is as follows.

THEOREM 5.1. *In order that the extremal g and set $(\alpha) = (0)$ afford a minimum to J relative to neighboring admissible curves and sets (α) it is sufficient that g and the set $(\alpha) = (0)$ satisfy the transversality condition, that there be no conjugate point of $x = a^1$ on the interval $a^1 < x < a^2$, and that the Clebsch, Weierstrass, and Mayer S-conditions hold.*

Hestenes [4] has shown that g and the set $(\alpha) = (0)$ afford the desired minimum provided the second variation is positive definite for admissible sets (η) and (u). Cf. Currier [5]. The problem here is accordingly to show that the second variation is positive for non-null admissible sets (η) and (u).

We observe that $x = a^2$ is not conjugate to $x = a^1$. For otherwise there would be a secondary extremal for which $\eta_i \not\equiv 0$, which would satisfy the end conditions with the the set $(u) = (0)$, and would appear with this $\eta_i(x)$ and set (u) as a member of the family (5.6) for which $(v) \neq (0)$. We would then have $H(v) = 0$, contrary to hypothesis.

It follows from the preceding sections that each secondary extremal for which $a^1 \leqq x \leqq a^2$ gives a proper minimum to I relative to admissible curves $\eta_i(x)$ which join its end points. That $I > 0$ for admissible secondary extremals and sets (u) for which (η) and (u) are not both null follows from the positive definiteness of $H(v)$. Hence I is positive definite as stated.

The proof of the theorem is complete.

BIBLIOGRAPHY

1. Bolza, *Über den "anormalen Fall" beim Lagrangeschen und Mayerschen Problem mit gemischten Bedingungen und variablen Endpunkten*, Mathematische Annalen, vol. 74 (1913), pp. 430–446.

2. Bliss, *The problem of Bolza in the calculus of variations*, Annals of Mathematics, vol. 33 (1932), pp. 261–274.

3. Carathéodory, *Üeber die Einteilung der Variationsprobleme von Lagrange nach Klassen*, Commentarii Mathematici Helvetici, vol. 5 (1933), pp. 1–19.

4. Hestenes, *Sufficient conditions for the problem of Bolza in the calculus of variations*, these Transactions, vol. 35 (1934).

5. Currier, *The variable end point problem of the calculus of variations including a generalization of the classical Jacobi conditions*, these Transactions, vol. 34 (1932), pp. 689–704.

6. Morse and Myers, *The problems of Lagrange and Mayer with variable end points*, Proceedings of the American Academy of Arts and Sciences, vol. 66 (1931), pp. 235–253.

7. Morse, *Sufficient conditions in the problem of Lagrange with fixed end points*, Annals of Mathematics, vol. 32 (1931), pp. 567–577.

8. Morse, *Sufficient conditions in the problem of Lagrange with variable end points*, American Journal of Mathematics, vol. 53 (1931), pp. 517–546.

9. Note added in proof. At the September meeting of the Society at Williamstown, Dr. Reid, unaware of the existence of the present paper, reported on a proof of theorems similar to the theorems contained herein. In the final theorems he assumed normalcy on the interval (a^1, a^2) as against the author's weaker assumption that a λ_0 exists which is not zero. More recently Dr. Hestenes has announced proofs of the theorems concerned.

HARVARD UNIVERSITY,
CAMBRIDGE, MASS.

METABELIAN GROUPS OF ORDER p^m, $p>2$*

BY

CHARLES HOPKINS

Introduction

A metabelian group is defined as one whose central quotient-group is abelian.† Since the central quotient-group of any group G is simply isomorphic with the group of inner isomorphisms of G, a metabelian group may also be defined as a group whose group of inner isomorphisms is abelian.‡

Any metabelian group G of order $p_1^{a_1}p_2^{a_2}\cdots p_n^{a_n}$ is the direct product of its Sylow subgroups of order $p_i^{a_i}$. In developing a theory of metabelian groups, it is accordingly reasonable to center the attention upon those of order p^m. In view of the fact that many results which are valid for groups of order p^m, $p>2$, do not hold for $p=2$, it seems advantageous to treat separately the cases $p=2$ and $p>2$. In this article we are concerned exclusively with the case $p>2$.

In §§2–5 we develop, by aid of the theory of regular permutation groups, certain general properties of a metabelian group G of order p^m, $p>2$. We mention the following:

(1) G is conformal with an abelian group A;

(2) the operations of any metabelian group which is conformal with A can be derived by making A isomorphic with a certain subgroup of its group of isomorphisms and multiplying together corresponding operations;

(3) the group of isomorphisms of G is a subgroup of the group of isomorphisms of A.

In §§7–9 we define four different types of bases for G and prove that each of these types occurs in every G. (Any set of elements which generate G is said to constitute a basis for G.) Two of these types, the MB-bases and the U-bases, are of fundamental importance in the theory of metabelian groups. In §§10–11 we exhibit certain relationships between these two types of bases and, furthermore, between the U-bases of G and those of A.

In §§12–14 we discuss the topic of abstract defining relations for G: with reference to a U-basis (§12), an MB-basis (§13), and a U-basis for A (§14).

---

* Presented to the Society, December 27, 1933; received by the editors March 22, 1934.

† W. B. Fite, Proceedings of the American Association for the Advancement of Science, vol. 49 (1901), p. 41.

‡ The term "metabelsche Gruppe," as used by Furtwängler and other German mathematicians, denotes a group whose commutator subgroup is abelian.

Two of the fundamental results of this paper—that G is conformal with an abelian group, and that G possesses a U-basis—have been published in a recent article by P. Hall, entitled *A contribution to the theory of groups of prime power order.** The author, however, feels it desirable to present his original proofs of these two results, as the methods involved are of frequent occurrence throughout this paper.

NOTATION, ELEMENTARY RESULTS

1. In order to avoid repeated explanations, the symbols employed in this article will usually preserve their significance throughout, and accordingly will ordinarily be defined only at their initial appearance.

The letter G will always denote a metabelian group of order p^m, $p>2$. The central and the commutator subgroup of G will be designated by Γ and C respectively. The operations of G will usually be denoted by small letters (s, σ etc.); for the automorphisms of G we shall always use capital letters.

The symbol c_{ij} shall denote the commutator $s_i^{-1}s_js_is_j^{-1}$ (or $\sigma_i^{-1}\sigma_j\sigma_i\sigma_j^{-1}$). Since each commutator is invariant in G, of the eight formally distinct commutators which arise from any two operations of G, only two, namely c_{ij} and c_{ji}, will be effectively distinct. Obviously c_{ji} equals c_{ij}^{-1}.

We now mention certain elementary results, which we shall tacitly assume throughout this paper.

(a) If $g_1, g_2, \cdots, g_n$ are any set of independent generating operations (I.G.O.) for G,† then C is generated by $c_{12}=g_1^{-1}g_2g_1g_2^{-1}, c_{13}, \cdots, c_{1n}, c_{23}, \cdots, c_{2n}, \cdots, c_{n-1,n}$.

(b) Every operation of G can be expressed in the form $g_1^{x_1}g_2^{x_2}\cdots g_n^{x_n} \cdot c_{12}^{x_{12}}c_{13}^{x_{13}}\cdots c_{n-1,n}^{x_{n-1,n}}$.

(c) If σ_a and σ_b are of orders p^{m_a} and p^{m_b} respectively, $m_a \geqq m_b$, then the order of $\sigma_a\sigma_b$ divides p^{m_a}.‡ From this we see that every set of I.G.O. for G must include operations of highest order in G.

(d) The product of any two pth powers in G is itself a pth power in G.§

DEFINITION. *Any operation of a given prime-power group which is not a pth power of an operation in this group is said to be a "principal element" of this given group.*

---

* Proceedings of the London Mathematical Society, (2), vol. 36, parts 1 and 2, pp. 29–95. The results presented in the author's paper were obtained prior to the appearance of Hall's article.

† That the number of elements in any set of I.G.O. for a prime power group is an invariant of the group was proved by G. A. Miller, these Transactions, vol. 16 (1915), p. 21.

‡ W. B. Fite, these Transactions, vol. 3 (1902), p. 338.

§ P. Hall, loc. cit., p. 75.

RESULTS DERIVED FROM THE REPRESENTATION OF G AS A REGULAR PERMUTATION GROUP

2. Let G denote any metabelian group of order p^m, $p>2$. Regarding G as an abstract group, we denote its operations by the symbols $\sigma_1, \sigma_2, \cdots, \sigma_i, \cdots$.

We denote any permutation s_i of G (in the regular representation derived from post-multiplication) by the symbol $\binom{\sigma}{\sigma\sigma_i}$. We may think of s_i as a representation of σ_i. A representation, as a permutation on the p^m symbols $\sigma_1, \sigma_2, \cdots$, of any inner isomorphism S_i of G is afforded by the symbol $\binom{\sigma}{\sigma_i^{-1}\sigma\sigma_i}$.* Clearly S_i transforms the operations of G according to the permutation s_i. The totality of distinct symbols S constitutes a representation H of the group of inner isomorphisms of G. Any element of $I(G)$, the group of isomorphisms of G, can be represented as a permutation on the letters of G by the symbol $\binom{\sigma}{\sigma'}$.

We may identify $I(G)$ with that subgroup of the holomorph $K(G)$ of G whose permutations omit the symbol for the identity of G.† Under G, $I(G)$ is transformed into p^d conjugates, where p^d equals p^m divided by the number of characteristic operations of G. The totality of distinct permutations in these p^d conjugates coincides with the totality of distinct products $\binom{\sigma}{\sigma_k\sigma}\binom{\sigma}{\sigma'}$, where $\binom{\sigma}{\sigma_k\sigma}$ is any permutation in the conjoint of G, while $\binom{\sigma}{\sigma'}$ is any permutation of $I(G)$.

Let p^μ denote the order of H. One readily sees that H has under G exactly p^μ conjugates.

Now $J\equiv\{G, H\}$ is a metabelian group of order $p^{m+\mu}$. Its central is Γ, its commutator subgroup is C; the central quotient-group J/Γ is the direct product of two simply-isomorphic groups, each of which is simply-isomorphic with H. The chief interest in J attaches to the fact that it contains a remarkable set of subgroups, each of which is conformal with G.

Since H is simply-isomorphic with G/Γ, we obtain an isomorphism of G with H by making each operation s of G correspond to that operation S of H which transforms G according to s. Let θ be defined as the operation of making G isomorphic with H in this manner and then multiplying together corresponding operations. That is, $\theta s = sS$. Similarly, we define θ_x by the equation $\theta_x s = sS^x$. (Since H is abelian, $s_1 \sim S_1^x$, $s_2 \sim S_2^x$, $\cdots$ etc. defines an isomorphism of G with H.) Let p^ν be the order of the operation of highest order in H. If

* This notation is fully explained in Speiser, *Theorie der Gruppen von endlicher Ordnung*, 2d edition, p. 25, p. 121; and in Burnside, *Theory of Groups*, 2d edition, pp. 81 ff.

† We agree that all the permutations of G shall begin with the symbol σ_1 for the identity of G. We may define $I(G)$ as that subgroup of $K(G)$ generated by its permutations which omit the initial letter in the permutations of G.

we let x range over all positive integral values, it is clear that not more than p^{ν} of the operations $\theta_x s$ will be distinct.

By the symbol $\theta_x G$ we designate the set of p^m operations which we obtain by applying to each operation of G the operator θ. We note that Γ is common to G and to $\theta_x G$. If we assume for the moment (we shall prove it below) that $\theta_x G$ is a group for which $\theta_x G/\Gamma$ is simply isomorphic with H, then we may define the symbol $\theta_y \theta_x$ by the equation $\theta_y \theta_x s = \theta_y(sS^x)S^y$. Then $\theta_y \theta_x = \theta_x \theta_y = \theta_{x+y}$. Hence we may write $\theta_x = \theta^x$, regarding θ as an operator of period p^{ν}.

THEOREM I. *The permutations $\theta_x G$ constitute a group which is conformal with G.*

That they constitute a group follows from the equation $(\theta^x s_i)(\theta^x s_j) = c_{ij}^{-x}\theta^x(s_i s_j) = \theta^x(c_{ij}^{-x} s_i s_j)$. That this group is conformal with G is evident from the fact that $\theta^x s$ and s have the same order. (For s and S are commutative and have only the identity in common, since s permutes all the letters of G while S omits σ_1. The order of S divides the order of s, since s and S transform the operations of G in the same way.)

THEOREM II. *Each G_x is an invariant subgroup of the holomorph of G.*

Clearly G is commutative with G_x, since G is commutative with Γ and transforms every coset $\Gamma s S^x$ into itself. To show that $I(G)$ transforms G_x into itself, we proceed as follows.

Let $\binom{\sigma}{\sigma'}$ denote any permutation of $I(G)$. The operation $s_i S_i^x$ of G_x may be represented as

$$\binom{\sigma}{\sigma\sigma_i}\binom{\sigma}{\sigma_i^{-x}\sigma\sigma_i^x}, \text{ which equals } \binom{\sigma}{\sigma_i^{-x}\sigma\sigma_i^{x+1}}.$$

Now

$$\binom{\sigma'}{\sigma}\binom{\sigma}{\sigma_i^{-x}\sigma\sigma_i^{x+1}}\binom{\sigma}{\sigma'} = \binom{\sigma'}{\sigma_i^{-x}\sigma\sigma_i^{x+1}}\binom{\sigma_i^{-x}\,\sigma\,\sigma_i^{x+1}}{\sigma_i'^{-x}\sigma'\sigma_i'^{x+1}} = \binom{\sigma'}{\sigma_i'^{-x}\sigma'\sigma_i'^{x+1}}$$

$$= \binom{\sigma}{\sigma_i'^{-x}\sigma\sigma_i'^{x+1}} = \binom{\sigma}{\sigma\sigma_i'}\binom{\sigma}{\sigma_i'^{-x}\sigma\sigma_i'^{x}} = s_i' S_i'^x,$$

where S_i' transforms the operations of G according to s_i'. Since s_i' is an operation of G, $s_i' S_i'^x$ is an operation of G_x. This demonstrates our theorem, since K is generated by G and $I(G)$.

THEOREM III. *Each G_x is a regular group.*

Since G is a regular group on the symbols $\sigma_1, \sigma_2, \cdots, \sigma_i, \cdots$, while every permutation of H omits σ_1, it is obvious that every permutation of G

other than the identity must permute σ_1. Suppose that some permutation t of G_x, distinct from the identity, omits the symbol σ_x. Now G contains a permutation $\bar{s}$ which replaces σ_k by σ_1. But $\bar{s}^{-1}t\bar{s}$ is a permutation of G_x which omits σ_1 (see Theorem II). This proves that each permutation of G_x other than the identity permutes all the symbols $\sigma_1, \sigma_2, \cdots, \sigma_{p^m}$. Since G_x is conformal with G, it must be a regular group on these symbols.

3. These p^ν conformal groups $G_1, G_2, \cdots, G_x, \cdots, G_{p^\nu}=G$ constitute a set which we shall refer to as D. As permutation groups in $K(G)$, they are all distinct. We shall prove shortly that regarded as abstract groups exactly $\nu+1$ of them are distinct.

Let $t_i=s_iS_i^x$ and $t_j=s_jS_j^x$ be any two operations of G_x. Now $t_i^{-1}t_jt_i=t_jc_{ij}^{2x+1}$, where $c_{ij}=s_i^{-1}s_js_is_j^{-1}$.* If $2x+1$ is prime to p, then the commutator subgroup of G_x coincides with the commutator subgroup C of G. If $2x+1$ is divisible by p but not by p^2, then the commutator subgroup of G_x is composed of the pth powers of the elements of C. By means of the relation $y\equiv 2x+1 \pmod{p^\nu}$, we may associate with each member of D a value of y as a subscript. The $p^{\nu-1}(p-1)$ members of D for which y is prime to p constitute a subset which we call D_1; the $p^{\nu-2}(p-1)$ members of D for which y is divisible by p, but not by p^2, we shall put into a set D_p, etc. Set D_{p^ν} consists of a single group, namely that G_x for which $2x+1$ is divisible by p^ν. This group, which is abelian, we shall designate by the letter A. Its permutations $t_1, t_2, \cdots, t_i, \cdots$ are connected with those of G by the equation $t_i=\theta^a s_i$, where a is the smallest positive root of $2a+1\equiv 0 \pmod{p^\nu}$. That G_a is abelian is sufficiently important to state as a theorem.

THEOREM I. *The p^m products $t_i=\theta^a s_i$, $i=1, 2, \cdots, p^m$, constitute a regular abelian group A which is conformal with G.*

The conjoint of each group in a given set D_{p^α}, $\alpha=0, 1, \cdots, \nu$, is a member of the same set. If y has the value k for a given group G_x, then y will be congruent to $-k$ modulo p^ν for the conjoint of G_x. (It is easy to prove that the conjoint of G_x is $G_{p^\nu-x-1}$.)

THEOREM II. *The groups in any given set D_{p^α}, $\alpha=0, 1, \cdots, \nu$, are simply isomorphic.*

Let λ be an integer prime to p. If we replace each operation of G_x by its λth power, then we shall obtain all the operations of G_x in some order. Let

$$T_\lambda = \begin{pmatrix} \sigma \\ \sigma^\lambda \end{pmatrix}$$

* It is a simple task to verify the relations $s_i^{-1}S_js_i=S_jc_{ij}$ and $S_i^{-1}s_jS_i=s_jc_{ij}$. Of course $S_i^{-1}S_jS_i=S_j$.

be the permutation on the symbols $\sigma_1, \sigma_2, \cdots$ derived from associating each operation of G with its λth power. Since T_λ defines an automorphism of the abelian group Γ, in determining how T_λ transforms the operations of G_x we shall be concerned only with the non-invariant operations of G_x.

Let $t_i = s_i S_i^x$ be any non-invariant operation of G_x. We may write

$$t_i = \begin{pmatrix} \sigma \\ \sigma_i^{-x}\sigma\sigma_i^{x+1} \end{pmatrix}.$$

Then

$$T_\lambda^{-1} t_i T_\lambda = \begin{pmatrix} \sigma^\lambda \\ \sigma \end{pmatrix}\begin{pmatrix} \sigma \\ \sigma_i^{-x}\sigma\sigma_i^{x+1} \end{pmatrix}\begin{pmatrix} \sigma \\ \sigma^\lambda \end{pmatrix} = \begin{pmatrix} \sigma^\lambda \\ [\sigma_i^{-x}\sigma\sigma_i^{x+1}]^\lambda \end{pmatrix}.$$

Now

$$[\sigma_i^{-x}\sigma\sigma_i^{x+1}]^\lambda = \sigma_i^\lambda[\sigma^{-x}\sigma\sigma_i^x]^\lambda c_{i\sigma}^{\lambda(\lambda+1)/2},$$

where $c_{i\sigma}$ is $\sigma_i^{-1}\sigma\sigma_i\sigma^{-1}$. Moreover,

$$[\sigma_i^{-x}\sigma\sigma_i^x]^\lambda = \sigma_i^{-x}\sigma^\lambda\sigma_i^x.$$

Hence

$$[\sigma_i^{-x}\sigma\sigma_i^{x+1}]^\lambda = \sigma_i^\lambda(\sigma_i^{-x}\sigma^\lambda\sigma_i^x)c_{i\sigma}^{\lambda(\lambda+1)/2}.$$

Since 2 is prime to p, the congruence $2z \equiv 1 \pmod{p^\nu}$ always admits a unique solution z. Therefore, we are justified in using the symbol $(\lambda+1)/2$, even when $\lambda+1$ is an odd integer. Now

$$\sigma^\lambda c_{i\sigma}^{\lambda(\lambda+1)/2} = \sigma_i^{-(\lambda+1)/2}\sigma^\lambda\sigma_i^{(\lambda+1)/2} = \sigma_i^{-(\lambda+1)/2}(\sigma_i^{-x}\sigma^\lambda\sigma_i^x)\sigma_i^{(\lambda+1)/2}.$$

Hence one may write

$$[\sigma_i^{-x}\sigma\sigma_i^{x+1}]^\lambda = \sigma_i^{-(1-\lambda)/2}(\sigma_i^{-x}\sigma^\lambda\sigma_i^x)\sigma_i^{(1-\lambda)/2+\lambda} = \sigma_i^{-(1-\lambda+2x)/2}\sigma^\lambda\sigma_i^{(1-\lambda+2x)/2}\sigma_i^\lambda.$$

Then

$$T_\lambda^{-1} t_i T_\lambda = \begin{pmatrix} \sigma \\ \sigma_i^{-(1-\lambda+2x)/2}\sigma\sigma_i^{(1-\lambda+2x)/2}\sigma_i^\lambda \end{pmatrix} = s_i^\lambda S_i^{(1-\lambda+2x)/2}.$$

Let us put $1-\lambda+2x \equiv 2\lambda\xi \pmod{p^\nu}$. Then $s_i^\lambda S_i^{\lambda\xi}$ is an operation of G_ξ. If we write the congruence above in the form $1+2x \equiv \lambda(1+2\xi) \pmod{p^\nu}$, it is clear that the same power of p divides both $1+2x$ and $1+2\xi$. This demonstrates our theorem. If G_x is the abelian group A, then $1+2x$ (and conse-

quently $1+2\xi$) is divisible by p^ν. We can, therefore, identify the permutation T_λ with that automorphism of A which transforms each operation of A into its λth power.

If we restrict the values of λ to the $p^{\nu-1}(p-1)$ positive integers which are less than p^ν and prime to p, then the permutations T_λ constitute a cyclic group of order $p^{\nu-1}(p-1)$. We may identify each T_λ with the linear substitution X_λ on the subscripts of $G_1, G_2, \cdots, G_x, \cdots$, where X_λ is $x' \equiv a[1-\lambda'(2x+1)]$ (mod p^ν), while a and λ' are defined by the congruences $2a+1\equiv 0$ (mod p^ν) and $\lambda\lambda'\equiv 1$ (mod p^ν). Or we can represent T_λ as the linear substitution $Y_\lambda: y'\equiv\lambda' y$ (mod p^ν). The order of T_λ is obviously the period of λ with respect to p^ν. When λ is $p^\nu-1$, then T_λ represents a substitution of order 2 in the double holomorph of G which transforms each G_x into its conjoint.

4. At this point we review certain results from §§2–3, which will be of service to us in what follows. Commencing with a representation of G as a regular permutation group, whose permutations $s_1, s_2, \cdots$ all begin with the same letter σ_1, we designate the holomorph of G (on these letters) by $K(G)$. We let $I(G)$ denote that representation in $K(G)$ of the group of isomorphisms of G whose permutations all omit σ_1. The subgroup of $I(G)$ which gives the inner isomorphisms of G we shall denote by H. We let $S_1, S_2, \cdots$ denote the permutations of H, where S_i transforms G according to s_i. Furthermore, p^ν denotes the order of the element of highest order in H, while a is the least positive root of $2a+1\equiv 0$ (mod p^ν). Theorem I of §3 states that (a) the p^m elements $t_i=\theta^a s_i = s_i S_i^a$ constitute a regular abelian group A on the letters of G. Let $K(A)$ denote the holomorph of A (on these same letters), and let $I(A)$ denote that representation in $K(A)$ of the group of isomorphisms of A whose permutations omit σ_1. Since $I(G)$ is a subgroup of $I(A)$,* H is in $I(A)$. Throughout the remainder of this article the symbols defined above will preserve their significance.

We know that there is only one permutation in $I(G)$ which transforms the permutations of G in a prescribed manner. Hence, given s_i in the equation $t_i=\theta^a s_i$, we see that t_i is uniquely determined. Conversely, given t_i in this equation, s_i is uniquely given by $t_i S_i^{-a}$. We recall that S_i^{-a} is in $I(A)$. We may, therefore, state that (b) the permutations in a given regular representation of G may be obtained from those of A by making A isomorphic with a certain subgroup of $I(A)$ and multiplying together corresponding operations. This result is clearly trivial in the sense that we cannot determine the "certain subgroup" unless we already know the permutations of G. The real point of (b) is expressed in the following theorem.

---

* See Theorem II of §2.

THEOREM I. *Let* $t_1, t_2, \cdots$ *denote the permutations of* A *and let* $R_1, R_2, \cdots$ *denote the permutations of a subgroup* $\overline{R}$ *of* $I(A)$. *Let* γ_{ij} *denote the commutator* $R_i^{-1}t_jR_it_j^{-1}$, *and* (c) *let every product* $\gamma_{ij}\gamma_{ji}^{-1}$ *be invariant under* $\overline{R}$. *If the correspondence* $\overline{\Gamma}\sim E, \cdots, \overline{\Gamma}t_i\sim R_i, \cdots$ *defines an isomorphism* τ *of* A *with* $\overline{R}$ *for which* (d) $\overline{\Gamma}$ *contains every* γ_{ij}, *then the* p^m *products*

$$\overline{\Gamma}E, \cdots, \overline{\Gamma}t_iR_i, \cdots \tag{1}$$

constitute a metabelian subgroup $\overline{G}$ *of* $K(A)$ *which is conformal with* A.*

From (d) we know that the product of any two elements in the set (1) is itself in the set. That the p^m products (1) are all distinct follows from the fact that A and $\overline{R}$ have only the identity in common. That these products constitute a metabelian group is a consequence of (c). From the existence of τ we know that the order of R_i divides the order of t_i. Although R_i and t_i are not necessarily commutative, a simple computation will show that t_iR_i and t_i have the same order. From this it will follow that $\overline{G}$ is conformal with A.†

To show that we obtain every regular metabelian group in $K(A)$ which is conformal with A by employing, in the procedure of Theorem I, every "permissible" subgroup $\overline{R}$ of $I(A)$, it is clearly sufficient to show that $\overline{G}$ is a regular permutation group. For we know that the permutations of a given regular permutation group G in $K(A)$ can be derived from those of A by the equation $s_i = t_iS_i^{-a}$. In §5 we shall prove that every $\overline{G}$ is a regular group.

That a representation as a regular permutation group of each of the abstractly distinct metabelian groups which are conformal with G may be obtained by the process of Theorem I, is a direct consequence of the following:

THEOREM II. *The holomorph* $K(A)$ *contains a regular representation of each of those abstractly distinct metabelian groups which are conformal with* A.

---

* The identical operation of any group is denoted by the letter E.

† We observe that t_i and R_i need not be commutative; moreover, the γ_{ij} need not be separately invariant under $\overline{G}$. But it is obvious that every commutator $R_k^{-1}\gamma_{ij}R_k\gamma_{ij}^{-1}$ must be invariant under $\overline{G}$, and hence under $\overline{R}$. That is, the class of $\{A, \overline{R}\}$ cannot exceed 2.

This derivation of $\overline{G}$ from A and $\overline{R}$ is a special example of a more general "composition" of two groups. We refer to the following theorem:

Let Q *and* Q' *be two finite groups of orders* m *and* m' *respectively, for which the following conditions hold*: (a) *the cross-cut of* Q *and* Q' *is the identity*; (b) Q' *transforms* Q *into itself*; (c) Q *and* Q' *are isomorphic under the correspondence* $\overline{Q}\sim E, \cdots, \overline{Q}q_i\sim q_i', \cdots$, *where* $\overline{Q}$ *contains all commutators* $q_j'^{-1}q_kq_j'q_k^{-1}$, q_k *and* q_j' *being any two elements of* Q *and* Q' *respectively. Then the* m *products*

$$\overline{Q}E, \cdots, \overline{Q}q_iq_i', \cdots \tag{1}$$

constitute a group Q'' *of order* m.

From (c) it is clear that $q_iq_i'q_jq_j'$ can be brought into the form $q_{ij}q_iq_jq_i'q_j'$, where q_{ij} is in $\overline{Q}$. That is, the product of any two elements in (1) is in the set (1). From (a) we see that these m products are distinct, since $q_iq_i' = q_jq_j'$ leads to $q_j^{-1}q_i = q_j'q_i'^{-1} = E$.

Of course, Q'' and Q are usually not conformal. The simplest additional restriction which will ensure their being conformal is probably that given by $q_iq_i' = q_i'q_i$.

Let G and G' be any two such groups, each being represented as a regular permutation group. Then $K(A)$ on the letters of G and $K(A)$ on the letters of G' are conjugate under some permutation. Hence G' occurs as a regular group in the holomorph $K(A)$ (on the letters of G).

5. In this section we develop several theorems which, in the main, are generalizations of theorems in §§2–3. The symbol $\overline{G}$ is the same as in Theorem I of §4.

THEOREM I. *Each $\overline{G}$ is transformed into itself by the permutations of A,* and conversely.*

We regard the elements of $\overline{G}$ as a certain p^m products t_iR_i (see Theorem I of §4). We write $\overline{G}$ in cosets with respect to Γ, where Γ is the subgroup of $\overline{G}$ (and of A) composed of those products for which R_i is the identity. Then the permutations of A transform each of these cosets into itself. This proves the first part of our theorem. The converse is obvious.

THEOREM II. *Each $\overline{G}$ is a regular group.*

As above, we regard the elements of G as the products t_iR_i. Since each permutation of $I(A)$ omits σ_1, the initial letter in the permutations of A, we see that every permutation in $\overline{G}$ permutes σ_1. If a certain permutation, say $\bar{t}$, of $\overline{G}$ should omit the letter σ_k, then we could find a permutation t in A such that $t^{-1}\bar{t}t$ would omit σ_1. From Theorem I of this section we know that this transform is in G. Hence each permutation of G permutes all the letters of A; G is accordingly regular, since it is conformal with A.

THEOREM III. *All simply-isomorphic regular metabelian groups G' in $K(A)$ which are conformal with A constitute a complete set of conjugates under $I(A)$.*

Let G' and G'' be any two of these simply-isomorphic regular groups in $K(A)$. Since G' and G'' are both regular, they are conjugate under some permutation on the letters of A. Our objective is to show that one such permutation occurs in $I(A)$.

We denote the group of inner isomorphisms of G' by H', and that of G'' by H''. Of course we regard H' and H'' as subgroups of $I(A)$. Let Γ' and Γ'' denote the centrals of G' and G'' respectively. Obviously Γ' and Γ'' are simply-isomorphic subgroups in A. We denote the permutations of G' by $s_1', s_2', \cdots$ and those of H' by $S_1', S_2', \cdots$, where S_i' transforms G' according to s_i'. We adopt a corresponding notation for G'' and H''.

To each permutation of A we assign two symbols, t' and t'', in such a

* Of course, we regard A as derived from a regular representation of a given metabelian group G.

way that the permutations of G' and A (G'' and A) are connected by the equation $t_i' = s_i' S_i'^{a}$ ($t_i'' = s_i'' S_i''^{a}$). We write $\bar{S}_i'$ for $S_i'^{-a}$ and $\bar{S}_i''$ for $S_i''^{-a}$. Then the permutations of G' and of G'' are derivable from those of A by the equations

$$s_i' = t_i' \bar{S}_i' \tag{1}$$

and

$$s_i'' = t_i'' S_i'' \tag{2}$$

respectively.

We may choose our notation so that a simple isomorphism between G' and G'' is defined by the correspondence

$$\Gamma' \sim \Gamma'', \cdots, \Gamma' t_i' \bar{S}_i' \sim \Gamma'' t_i'' \bar{S}_i'', \cdots, \Gamma' t_j' \bar{S}_j' \sim \Gamma'' t_j'' \bar{S}_j'', \cdots . \tag{3}$$

Now (3) requires that H' and H'' be isomorphic under the correspondence

$$\cdots, \bar{S}_i' \sim \bar{S}_i'', \cdots, \bar{S}_j' \sim \bar{S}_j'', \cdots . \tag{4}$$

Since the product $s_i' s_j'$ corresponds to $s_i'' s_j''$, we obtain $\gamma_{ij}' t_i' t_j' \bar{S}_i' \bar{S}_j' \sim \gamma_{ij}'' \cdot t_i'' t_j'' \bar{S}_i'' \bar{S}_j''$, where $\gamma_{ij}' = \bar{S}_i' t_j' \bar{S}_i'^{-1} t_j'^{-1}$ and $\gamma_{ij}'' = \bar{S}_i'' t_j'' \bar{S}_i''^{-1} t_j''^{-1}$. Since γ_{ij}' must correspond to γ_{ij}'' under (3), we get

$$(t_i' t_j')(\bar{S}_i' \bar{S}_j') \sim (t_i'' t_j'')(\bar{S}_i'' \bar{S}_j''). \tag{5}$$

From (4) and (5) we see that (3) involves an automorphism of A, defined by the correspondence

$$\Gamma' \sim \Gamma'', \cdots, \Gamma' t_i' \sim \Gamma'' t_i'', \cdots, \Gamma' t_j' \sim \Gamma'' t_j'', \cdots . \tag{6}$$

Let $\bar{\Pi}$ be the permutation in $I(A)$ which brings about the automorphism (6). Now $\bar{\Pi}$ transforms G' into a simply-isomorphic group $\bar{\Pi}^{-1} G' \bar{\Pi}$, and one readily sees that the permutations of these two groups correspond according to

$$\Gamma' \sim \Gamma'', \cdots, \Gamma' t_i' \bar{S}_i' \sim \Gamma'' t_i'' \bar{\Pi}^{-1} \bar{S}_i' \bar{\Pi}, \cdots . \tag{7}$$

Since $\bar{\Pi}^{-1} \gamma_{ij}' \bar{\Pi}$ equals γ_{ij}'', it follows that $\bar{\Pi}^{-1} \bar{S}_i' \bar{\Pi}$ and $\bar{S}_i''$ transform the permutations of A in the same way. But there is only one permutation in $I(A)$ which transforms A in a prescribed manner. Hence $\bar{S}_i''$ is $\bar{\Pi}^{-1} \bar{S}_i' \bar{\Pi}$ and G'' coincides with $\bar{\Pi}^{-1} G' \bar{\Pi}$. This completes the demonstration of Theorem III.

As a direct consequence of Theorem III we have

THEOREM IV. *A representation $I(G')$ of the group of isomorphisms of every G' occurs as a subgroup of $I(A)$.*

From Theorem III and Theorem IV follows

THEOREM V. *The number of distinct representations of G' in $K(A)$ equals the index of $I(G')$ in $I(A)$.*

THEOREM VI. *Those conjugates of G' (under $I(A)$) which are in the holomorph $K(G')$ of G' are commutative, each with each, and conversely.*

The proof is elementary. Equally obvious is

THEOREM VII. *The holomorph $K(G')$ is invariant in $K(A)$ if, and only if, the commutator subgroup of G' is a characteristic subgroup of A.*

In retrospect: Theorem I of §4 provides, theoretically at least, a means of constructing a regular representation of each of the abstractly distinct metabelian groups which are conformal with a given one G. By using every subgroup $\overline{R}$ of $I(A)$ which satisfies the conditions laid down in this theorem, we obtain the totality of regular metabelian groups in $K(A)$ which are conformal with A. Obviously we obtain this same totality of groups by subjecting the elements of $\overline{R}$ to the following additional restrictions: each γ_{ij} is invariant under $\overline{R}$; $\gamma_{ij}=\gamma_{ji}^{-1}$ (whence follows $\gamma_{ii}=E$).

The process of Theorem I in §4 does not, in general, yield all the metabelian groups in $K(A)$ which are conformal with A. In fact, when A has more than 2 I.G.O., then $K(A)$ always contains non-regular metabelian groups which are conformal with A. We shall not prove this statement; the demonstration is fairly obvious. Instead, we present the following example.

Let A be a representation as a regular permutation group of the abelian group of order 27 and type 1, 1, 1. We begin all the permutations of A with letter a_1; we denote by $I(A)$ the subgroup of $K(A)$ composed of the permutations of $K(A)$ which omit a_1. Now we can find in A three permutations A_1, A_2, A_3 which generate A. Also, we can find in $I(A)$ a permutation π of order 3 which is commutative with A_1 and A_3 and transforms A_2 into A_2A_1. We see that π permutes exactly 18 letters. Since π transforms $\{A_1, A_2\}$ into itself, it follows that $\{A_1, A_2, \pi\}$ is a metabelian group of order 27, each of whose operations is of order 3. This group is clearly not a regular group. It is of course simply isomorphic with the regular permutation group $\overline{G}=\{A_1, A_2, A_3\pi\}$, since all metabelian groups conformal with this given A are abstractly identical.

ARITHMETICAL INVARIANTS OF G

6. Associated with every given metabelian group G are the following uniquely-determined abelian groups: A, C, Γ, G/C, G/Γ (which is simply isomorphic with H). From each of these groups there arises a set of arith-

metical invariants of G. We enumerate the following:

(1) the r invariants $p^{\delta_1}, p^{\delta_2}, \cdots, p^{\delta_r}$ of A;
(2) the l invariants $p^{q_1}, \cdots, p^{q_l}$ of C;
(3) the invariants $p^{\nu}, p^{\nu}, p^{\nu_3}, \cdots, p^{\nu_h}$ of G/Γ;
(4) the invariants of Γ;
(5) the invariants of G/C.

To these we add

(6) the number n of I.G.O. for G. We may assume that for each set the invariants are arranged in descending order of magnitude.

We shall not go into the question of relationships between these six invariants other than to note that q_1 equals ν, while G/Γ must clearly have at least two invariants of highest order.

There are two additional invariants of G which are of considerable importance for the development of our theory. These we now proceed to define. Following the notation of Hall, we denote by $\mho_\alpha$ the subgroup composed of the p^αth powers of the elements of G. These groups $\mho_\alpha$ constitute a series of characteristic subgroups $\mho_1, \mho_2, \cdots, \mho_{\delta_1} = E$ of G, each being contained in those which precede it.* For C we define the series $C_1, C_2, \cdots, C_\nu = E$, where C_α is the subgroup composed of the p^αth powers of the elements of C. Finally we have a third series $\overline{C}_1, \overline{C}_2, \cdots, \overline{C}_\nu = E$, where $\overline{C}_\alpha$ is the subgroup composed of those elements of C which are in $\mho_\alpha$. Obviously $\overline{C}_\alpha$ contains C_α. In what follows we shall be concerned exclusively with the case $\alpha = 1$, i.e. with the first term in each of these three series.

The two additional invariants of G, to which we referred above, are the following:

(7) the number l_1 of invariants of $C/\overline{C}_1$;
(8) the number l_2 of invariants of $\overline{C}_1/C_1$.†

Since $C/\overline{C}_1$ and $C/C_1 \div \overline{C}_1/C_1$ are simply-isomorphic, we have $l_1 + l_2 = l$.

For the case where $C/\overline{C}_1$ is the identity there arise so many important simplifications of the general theory that it is desirable to assign a name to those groups G for which $C \equiv \overline{C}_1$. Such groups we shall call *ω-groups*. An immediate illustration of their significance is provided by

THEOREM I. *The quotient group $G/\mho_1(G)$ is abelian if, and only if, G is an ω-group.*

The proof follows from the fact that $C/\overline{C}_1$ is the commutator subgroup of

* Hall, loc. cit., p. 78.

† Obviously the two quotient-groups $C/\overline{C}_1$ and $\overline{C}_1/C_1$ are of type $1, 1, \cdots$.

$G/\mho_1(G)$.*

For certain small values of m (for $m=3$ and $m=4$, in particular) these eight invariants characterize G. It would be an interesting problem to determine whether for every order of G there exists a set of arithmetical invariants which completely characterize G: that is, determine G to within an isomorphism.

At this point we mention several useful properties of the ϕ-subgroup $\Phi(G)$ of G. That (a) $\Phi(G)$ is the cross-cut of all subgroups of index p in G, and that (b) G/Φ is of order p^n and type $1, 1, \cdots, 1$, are two familiar results in the theory of prime-power groups. From (b) it follows that $\Phi(G)$ is generated by $\mho_1(G)$ and C.

Now $\Phi(A)$, the ϕ-subgroup of A, coincides with $\mho_1(A)$. Obviously $\mho_1(A)$ is $\theta^a\mho_1(G)$. Since $\theta^a C$ is C itself, we have

THEOREM II. *The quotient-group $\theta^a\Phi(G)/\Phi(A)$ coincides with $C/\overline{C}_1$, and is accordingly of type* $1, 1, \cdots, 1$ *to l_1 factors.*

From Theorem II follows

THEOREM III. *For G to be an ω-group it is necessary and sufficient that r equal n. If G is not an ω-group, then $r-n$ must equal l_1.*

We mention here two rather obvious results, which will be of use to us in what follows. The first is the following: if $g_1, \cdots, g_n$ are a set of I.G.O. for G, then no product $g_1^{x_1}g_2^{x_2}\cdots g_n^{x_n}$ in which an exponent is prime to p can be in $\Phi(G)$. The second is

THEOREM IV. *If G is an ω-group, then every operation of G can be expressed in the form $g_x=g_1^{x_1}g_2^{x_2}\cdots g_n^{x_n}$.*

To prove this, we note that every operation σ in G can be expressed in the form g_yc, where g_y is $g_1^{y_1}g_2^{y_2}\cdots g_n^{y_n}$ and c is some element in C. Since C is in $\mho_1(G)$, c can be expressed as $(g_zc')^p$. Then σ can be brought into the form $g_yg_z^pc'^p=g_{y'}c''$. Since the order of c'' is less than the order of c, we can eventually bring σ into the form g_x above.

BASES FOR G

7. Any set of elements $g_1, g_2, \cdots$ which generate G we shall call a basis for G. In the classical theory of abelian groups the term *basis for Q*, where Q is an abelian group, is used to designate a set of elements $q_1, q_2, \cdots$ of Q such that Q is the direct product of the cyclic subgroups $\{q_1\}, \{q_2\}, \cdots$. To

* From a theorem of Hall (loc. cit., p. 83) it results that $\mho_\alpha/\mho_{\alpha+1}$ is abelian for $\alpha>1$. More generally, $\mho_\alpha/\mho_{\alpha+\beta}$ is abelian for $\alpha>1$ and $\alpha\geqq\beta$.

avoid confusion, a basis for Q of this sort we shall refer to as a U-basis for Q. The fact that every abelian group has a U-basis constitutes the so-called fundamental theorem in the theory of abelian groups.

We now define four different types of bases for G. For the first three types we start with a set of I.G.O. for G, namely $g_1, g_2, \cdots, g_n$. Let g_x be any product $g_1^{x_1}g_2^{x_2}\cdots g_n^{x_n}$ in which the g_i occur, without repetitions, in the normal order $g_1, g_2, \cdots$ etc.*

(1) If g_x is an element in C only when each factor $g_i^{x_i}$ is in C, then $g_1, \cdots, g_n$ are said to constitute a *C-basis* for G.

(2) If g_x is in Γ only when each factor $g_i^{x_i}$ is in Γ, then the g's are said to constitute a *Γ-basis* for G.

(3) If g_x is the identity only when each $g_i^{x_i}$ is the identity, the g's are said to constitute a *B-basis* for G.

A set of elements $P_1, P_2, \cdots, P_\rho$ is said to constitute a *uniqueness-basis* (*U-basis*) for G provided that each operation of G can be represented uniquely in the form $P_1^{x_1}P_2^{x_2}\cdots P_\rho^{x_\rho}$, where each exponent is a least positive residue modulo the order of the P to which it belongs.

In what follows we shall prove that each of these four types of bases occurs in any given metabelian group G. To prove the existence of a C-basis is a simple task. We write G in cosets with respect to C and select from every coset which corresponds to an element of a given U-basis for G/C an operation v_i. Now $v_1^{x_1}v_2^{x_2}\cdots$ is clearly in C only if each $v_1^{x_1}$, $v_2^{x_2}$, etc., is in C. It remains only to show that these v's constitute a set of I.G.O. for G. That they do is readily apparent from the relation $G/\Phi(G)\equiv G/C\div\Phi(G)/C$.

To construct a Γ-basis for G we first write G in cosets with respect to Γ; then, from each of those cosets which correspond to the elements of a U-basis for G/Γ, we select an operation of G, obtaining thereby the h operations $u_1, u_2, \cdots, u_h$.

If $h=n$, then $u_1, u_2, \cdots, u_h$ will constitute a set of I.G.O. for G. For $u_x=u_1^{x_1}u_2^{x_2}\cdots u_h^{x_h}$ is non-invariant in G unless each x_i is divisible by p^{ν_i}; we know that $\Phi(G)$ is $\{\mho_1(G), C\}$; hence u_x cannot be in $\Phi(G)$ unless each x is divisible by p.

If h is less than n, we can extend $u_1, \cdots, u_h$ to a set of I.G.O. for G by adding a certain $n-h$ elements $u_{h+1}, \cdots, u_n$. To show that these $n-h$ elements may be chosen from Γ, we observe that $G/\{u\}$ and $\Gamma/\overline{\Gamma}$ are simply-isomorphic, where $\{u\}$ is the group generated by $u_1, \cdots, u_h$, and $\overline{\Gamma}$ is the cross-cut of $\{u\}$ and Γ. Consequently $G\equiv\{\Gamma, \{u\}\}$; hence $u_{h+1}, \cdots, u_n$ may be taken from Γ.

* Throughout this paper it is assumed that in any indicated product, such as $A_1^{x_1}\cdots A_r^{x_r}$, $P_1^{x_1}\cdots P_r^{x_r}$ etc. no subscript is repeated.

We shall make no further use of these two types of bases. It is, perhaps, worthwhile to point out that for a given G it is usually impossible to construct a basis whose elements satisfy any two of the conditions (1), (2), (3) above. But there are large and important categories of groups G for which every C-basis is a U-basis. We mention, in particular, those groups G for which $h=n$ and G/Γ is of type $\alpha, \alpha, \cdots, \alpha$. Those groups G in which every element (except the identity) is of order p provide a trivial illustration of the case where every C-basis is simultaneously a Γ-basis and a B-basis.

8. We shall now prove that every metabelian group G contains a B-basis. That is, we shall show that there always exists a set of n I.G.O., $\beta_1, \beta_2, \cdots, \beta_n$ with the property that $\beta_1^{\lambda_1}\beta_2^{\lambda_2} \cdots \beta_n^{\lambda_n}$ is the identity only when each $\beta_i^{\lambda_i}$ is the identity.*

We first prove the theorem for groups having two I.G.O. Every set of I.G.O. must include at least one operation of highest order in G. Let β_1 be such an operation, and let β_2 be an operation in G of lowest possible order such that β_1 and β_2 generate G. We shall now prove that $\{\beta_1\}$ and $\{\beta_2\}$ can have only the identity in common.

Suppose that $\beta_2^{-p^{e_2}}=\beta_1^{bp^{e_1}}$, where $\beta_2^{-p^{e_2}}$ is not the identity. Since β_1 is of highest order in G, we may put $e_1=e_2+e_3$, where $e_3 \geqq 0$. Now

$$[\beta_2\beta_1^{bp^{e_3}}]^{p^{e_2}} = \beta_1^{bp^{e_1}}\beta_2^{p^{e_2}}[\beta_2\beta_1^{bp^{e_3}}\beta_2^{-1}\beta_1^{-bp^{e_3}}]^{p^{e_2}(p^{e_2}+1)/2}.$$

(Since p is an odd prime, $(p^{e_2}+1)/2$ is an integer; since $\beta_2^{p^{e_2}}$ is commutative with β_1, the order of the element in the brackets divides p^{e_2}.) Clearly β_1 and $\beta_2\beta_1^{bp^{e_3}}$ generate G. But the order of this second operation is less than the order of β_2, contrary to assumption. Our theorem, then, is true when G has two I.G.O.

We proceed by induction, assuming the validity of the theorem for all groups which have less than n I.G.O. Suppose, now, that G has n I.G.O. Among the operations of G which can occur in a set of I.G.O., let $\bar{s}$ be one of the smallest possible order. We consider the totality of sets of I.G.O. in which $\bar{s}$ occurs. For any one of these sets, say $\bar{s}_1, \bar{s}_2, \cdots, \bar{s}_{n-1}, \bar{s}$ the first $n-1$ elements generate a metabelian (or abelian) subgroup $\overline{H}$ of G. Since G is $\{\overline{H}, \bar{s}\}$, it is clear that $\overline{H}$ has exactly $n-1$ I.G.O. Hence for $\overline{H}$ we can find a B-basis, say $\beta_1, \beta_2, \cdots, \beta_{n-1}$. We now show that $\beta_1, \cdots, \beta_{n-1}, \bar{s}=\beta_n$ must constitute a B-basis for G. Let us assume the contrary; that is, let $\beta_1^{\lambda_1}\beta_2^{\lambda_2} \cdots \cdot \beta_{n-1}^{\lambda_{n-1}}\beta_n^{\lambda_n}=E$, where at least one of the λ's is not divisible by the order of the β to which it belongs. Certainly one of these λ's must be λ_n, since $\beta_1, \cdots, \beta_{n-1}$

* This result, in a slightly different form, was proved earlier by the author: Annals of Mathematics, vol. 29 (1928), pp. 6–9.

are a B-basis for $\overline{H}$. We have, then, $\beta_n^{\lambda_n}=\bar{\beta}$, where $\bar{\beta}$ is some element of $\overline{H}$. If G contains an element σ such that $\sigma^{\lambda_n}=\beta_n^{-\lambda_n}$, then $\sigma\beta_n$ will be of order λ_n (which is less than the order of $\bar{s}$). Since G is $\{\beta_1, \cdots, \beta_{n-1}, \sigma\beta_n\}$, this will contradict our assumption concerning $\bar{s}$. From this assumption we know that the order of each of the elements $\beta_1, \cdots, \beta_{n-1}$ is at least equal to the order of $\bar{s}$. Hence each constituent $\beta_i^{\lambda_i}$ in $\bar{\beta}$ can be regarded as $[\beta_i^{b_i}]^{\lambda_n}$.

Now $[\beta_1^{b_1}\beta_2^{b_2} \cdots \beta_{n-1}^{b_{n-1}}]^{\lambda_n}$ can be brought into the form $\beta_1^{\lambda_1} \cdots \beta_n^{\lambda_{n-1}}c^{k\lambda_n}$, where c is some element in the commutator subgroup of $\overline{H}$. If $c^{k\lambda_n}$ is the identity, then $\beta_1^{b_1} \cdots \beta_{n-1}^{b_{n-1}}$ will serve as the operation σ. If not, then we can find an element c' in the commutator subgroup of G such that $\beta_1^{b_1} \cdots \beta_{n-1}^{b_{n-1}} c'^k$ raised to the power λ_n will equal $\beta_n^{-\lambda_n}$.* We are led, then, to the conclusion that no relation $\beta_1^{\lambda_1}\beta_2^{\lambda_2} \cdots \beta_n^{\lambda_n}=E$ can exist unless each λ is divisible by the order of its β. This demonstrates our theorem.

One naturally asks whether for a certain permutation of the subscripts $1, \cdots, n$ there can exist a relation $\beta_{r_1}^{\mu_1}\beta_{r_2}^{\mu_2} \cdots \beta_{r_n}^{\mu_n}=E$, where at least one exponent is less than the order of its β. This we can answer in the negative. Such a relation could be brought into the form

$$(\beta_1^{\lambda_1} \cdots \beta_n^{\lambda_n})(c_{12}^{k_{12}\lambda_1\lambda_2} \cdots c_{n-1,n}^{k_{n-1,n}\lambda_{n-1}\lambda_n}) = (\beta)(c) = E.$$

Since (β) could not be the identity, we should have $(\beta)=(c)^{-1}\neq E$. Since (β) and (c) would be of the same order, at least two of the λ's would necessarily be prime to p. But (β) could then occur in a set of I.G.O. for G, while (c) obviously cannot have this property.

9. That every metabelian group G possesses a U-basis follows from the recent work of Hall in the field of prime-power groups.† The author, however, wishes to present his original proof, since the details are widely applicable in the following sections.

From the definition, it is clear that either (A) and (B) or (A) and (C) below provide a set of necessary and sufficient conditions for the elements $P_1, P_2, \cdots, P_\rho$ to constitute a U-basis for G:

(A) $P_x=P_1^{x_1}P_2^{x_2} \cdots P_\rho^{x_\rho}$ and $P_y=P_1^{y_1} \cdots P_\rho^{y_\rho}$ represent the same operation of G only when each x_i-y_i is divisible by the order of P_i;

(B) the product of the orders of $P_1, P_2, \cdots, P_\rho$ equals the order of G;

(C) every operation of G is representable in the form P_x.

Although it is not essential, we shall nevertheless find it convenient to

* Cf. (d) of §1.

† Hall, loc. cit., pp. 90–95. Hall proved (a) that every *regular* p-group is conformal with an abelian group; (b) that every regular p-group has a U-basis; (c) that the orders of the elements in a U-basis are the invariants of the conformal abelian group.

regard G as a regular permutation group. We shall, therefore, assume that the symbols which we employ have the meanings given in the first paragraph of §4. For the permutation of H which transforms G according to P_i we shall use the letter S_i. For convenience, we shall usually write T_i in place of S_i^{-a}.

Before demonstrating the existence of a U-basis for G, we shall prove the following result:

THEOREM I. *If $P_1, P_2, \cdots, P_\rho$ constitute a U-basis for G, then $P_1S_1^a$, $P_2S_2^a, \cdots, P_\rho S_\rho^a$ will constitute a U-basis for A.*

For the elements $A_1, A_2, \cdots$ to constitute a U-basis for A the following conditions are clearly sufficient:

(i) the product of the orders of $A_1, A_2, \cdots$ equals the order of A;

(ii) the product $A_x = A_1^{x_1}A_2^{x_2}\cdots$ can be the identity only when each factor $A_1^{x_1}, A_2^{x_2}, \cdots$ is the identity.

Now the orders of P_i and $P_iS_i^a$ are the same. Since the product of the orders of $P_1, P_2, \cdots$ must equal p^m, we see that (i) is satisfied for the elements $P_iS_i^a$.

We proceed to show that (ii) is also satisfied. We know that $P_x = P_y$ requires that $x_i - y_i$ be divisible by the order of P_i. Now $P_x = P_y$ is equivalent to $P_xP_y^{-1} = E$, and this latter equation may be brought into the form

(1) $P_xP_y^{-1} = P_1^{x_1-y_1}\cdots P_\rho^{x_\rho-y_\rho}\prod_{i<j} c_{ij}^{-y_i(x_j-y_j)} = E,$* where $c_{ij} = P_i^{-1}P_j^{-1}P_iP_j^{-1}$.

In (1) we write z_i in place of $x_i - y_i$, obtaining

$$P_{y+z}P_y^{-1} = P_1^{z_1}P_2^{z_2}\cdots P_\rho^{z_\rho}\prod c_{ij}^{-y_iz_j} = E. \tag{2}$$

Now the product

$$A_u = (P_1S_1^a)^{u_1}(P_2S_2^a)^{u_2}\cdots(P_\rho S_\rho^a)^{u_\rho}$$

can be reduced to the form

$$A_u = P_1^{u_1}\cdots P_\rho^{u_\rho}\prod c_{ij}^{-au_iu_j}\prod S_i^{au_i}. \tag{3}$$

Let us suppose that A_u is the identity. Since G and H are isomorphic under the correspondence $P_i \sim S_i^a$, we see that $A_u = E$ requires $\Pi S_i^{au_i} = E$.

Since the P_i are a U-basis for G, we know that equation (2) holds only when z_i is divisible by the order of P_i. By taking y_i equal to au_i and z_j equal to u_j, we see from (3) that A_u can be the identity only when the order of $P_iS_i^a$ divides u_i. This completes our proof.

* Since the elements c_{ij} are commutative, we may use the product sign Π. It is nevertheless desirable to think of the subscripts as occurring in a definite order, preferably the order $12, 13, \cdots, 1\rho, 23, \cdots, 2\rho, \cdots, \rho-1\ \rho$.

Since the orders of the elements in any U-basis for A are an invariant of A, we have, as a corollary,

THEOREM II. *The orders of the elements in any U-basis for G are the invariants $p^{\delta_1}, p^{\delta_2}, \cdots, p^{\delta_r}$ of A.**

We now state two theorems which assert the existence of a U-basis for any G.

THEOREM III. *If G is an ω-group and $A_1, \cdots, A_r$ are the elements of any U-basis for A, then $A_1T_1, A_2T_2, \cdots, A_rT_r$ will constitute a U-basis for G.*

THEOREM IV. *If G is any metabelian group of order p^m, $p>2$, and $A_1, \cdots, A_r$ constitute a primary U-basis for A, then a U-basis for G is given by the elements $A_1T_1, A_2T_2, \cdots, A_rT_r$.*

First we state what is meant by the term *primary U-basis for A*.

The U-basis $A_1, \cdots, A_r$ is said to be a *primary U-basis for A* provided that for the associated automorphisms $T_1, \cdots, T_r$ (arising from the equation $s_i=\theta^{-a}A_i=A_iT_i$) any product $T_x=T_1^{x_1}T_2^{x_2}\cdots T_r^{x_r}$ can be the identity only when the exponent of each T_i which is a principal element of H is divisible by p. For the present we shall assume that A possesses at least one primary U-basis; the proof will be given in the following section.

If $A_1, \cdots, A_r$ are any given U-basis for A, then the elements $A_1T_1, \cdots, A_rT_r$ have the property mentioned in (B) above. In determining whether the A_iT_i constitute a U-basis for G, the investigation, therefore, centers upon the equation

$$(A_1T_1)^{x_1}(A_2T_2)^{x_2}\cdots(A_rT_r)^{x_r}=(A_1T_1)^{y_1}\cdots(A_rT_r)^{y_r}. \tag{4}$$

This equation we may bring into the form

$$A_{x-y}\gamma_{x-y}T_{x-y}=E, \tag{5}$$

where

$$A_{x-y}=\prod_{i=1}^{r}A_i^{x_i-y_i},\quad \gamma_{x-y}=\prod_{i<j}\gamma_{ij}^{(x_i+y_i)(x_j-y_j)},$$

γ_{ij} being

$$T_iA_jT_i^{-1}A_j^{-1};\quad T_{x-y}=\prod_{1}^{r}T_i^{x_i-y_i}.$$

For convenience we shall write z_i in place of x_i-y_i and u_i in place of x_i+y_i. One easily sees that equation (5) requires $T_z=E$. Hence (5) reduces to

* Cf. Hall, loc. cit., p. 90.

$$A_z\gamma_z = E. \tag{6}$$

If (6) is satisfied only by $A_z = E$, then the A_iT_i will constitute a U-basis for G, since $A_z = E$ requires $z_i \equiv 0 \pmod{p^{\delta_i}}$. (The A_i constitute a U-basis for A.) Our objective is to show that when the A_i are selected according to the hypothesis of Theorem III or of Theorem IV, then (6) can be satisfied only by $A_z = E$.

(i) We assume that there exist certain values for the z_i and u_j such that A_z equals γ_z^{-1}, where A_z is not the identity. Then A_z and γ_z must be of the same order.

(ia) If G is an ω-group, then C must be a subgroup of $\mho_1(A)$. We observe that each commutator $\gamma_{ij}^{u_iz_j}$ in γ_z arises from T_i^u and the constituent $A_j^{z_j}$ of A_z. Since no element of a U-basis for A can be in $\mho_1(A)$, one readily sees that A_z and γ_z cannot be of the same order. For an ω-group, therefore, any U-basis of A leads to a U-basis for G.

(ib) Suppose that G is not an ω-group. We now assume that $A_1, \cdots, A_r$ are the elements of a primary U-basis for A. We wish to show that the assumption

$$A_z = \gamma_z^{-1},\ A_z \neq E, \tag{7}$$

is an impossible one.

As an element of A, each γ_{ij} can be expressed in the form $A_1^{b_1}A_2^{b_2}\cdots A_r^{b_r}$. If the exponent of every γ_{ij} in γ_z is divisible by p^α (but not by $p^{\alpha+1}$), then each exponent z_i in A_z must be divisible by p^α. In this case there must exist an element $\overline{A}_\alpha$ in A, whose order does not exceed p^α, such that $\overline{A}_\alpha A_{z'}$ equals $\gamma_{z'}^{-1}$, where z_i is $p^\alpha z_i'$, while $A_{z'}$ and $\gamma_{z'}$ are derived from A_z and γ_z respectively by substituting z_i' for z_i, leaving u_j unchanged. Then at least one of the exponents in $\gamma_{z'}$ will be prime to p. As we shall see, the argument is unaffected by the presence of the factor $\overline{A}_\alpha$, since $\overline{A}_\alpha$ is of lower order than $A_{z'}$. We shall, therefore, assume that in equation (7) the exponent of one of the γ_{ij}, say of $\gamma_{ab}^{u_az_b}$, is prime to p. Then $A_b^{z_b}$ must be a principal element of A. Since $A_b^{z_b}$ occurs in γ_z^{-1}, some constituent of γ_z, say $\gamma_{cd}^{u_cz_d}$, must contain A_b^λ, where λ is some exponent prime to p. Obviously u_cz_d must be prime to p, and γ_{cd} must be a principal element of A. Consequently T_d must be a principal element of H.

We recall that equation (5) is possible only when

$$T_z = T_1^{z_1}T_2^{z_2}\cdots T_d^{z_d}\cdots T_r^{z_r} \text{ is the identity.}$$

But the assumption that $A_1, \cdots, A_r$ are a primary U-basis and the conclusion above that z_d must be prime to p are clearly incompatible. In the case

of a primary U-basis $A_1, \cdots, A_r$, the assumption (i) can never be realized. This completes our demonstration of Theorem IV.

10. Theorem IV of §9 is clearly of little value unless we prove that A contains a primary U-basis. We indicate a method for constructing a primary U-basis, starting with any U-basis $A_1, \cdots, A_r$ of A. The order of A_i is of course p^{δ_i}; we assume the inequalities $\delta_1 \geqq \delta_2 \geqq \cdots \geqq \delta_r$.

It is a well known fact that the r elements

$$A_i' = A_1^{a_{i1}} A_2^{a_{i2}} \cdots A_r^{a_{ir}} \qquad (i = 1, 2, \cdots, r) \tag{1}$$

will constitute a U-basis for A, provided that the a_{ij} are any integers for which (a) the determinant $|a_{ij}|$ is prime to p, and (b) a_{ij} is divisible by $p^{\delta_j - \delta_i}$ for $i>j$. We propose to determine the a_{ij} so that $A_1', \cdots, A_r'$ will be a primary U-basis for A.

If the T_i's satisfy no relation of the form

$$T_1^{\lambda_1} T_2^{\lambda_2} \cdots T_r^{\lambda_r} = E \tag{2}$$

in which a λ is prime to p, then the initial U-basis $A_1, \cdots, A_r$ will be a primary U-basis. In the contrary case, let λ_α be the first λ in the sequence $\lambda_1, \lambda_2, \cdots$ which is prime to p, taking into account the totality of relations of type (2). If T_α is the identity, we eliminate T_α from every relation of type (2) and proceed to the next λ which is prime to p. If not, we replace A_α (in the set $A_1, \cdots, A_r$) by

$$A_\alpha' = A_\alpha^{\lambda_\alpha} A_{\alpha+1}^{\lambda_{\alpha+1}} \cdots A_r^{\lambda_r}.$$

Then for the permutation T_α' of H which is associated with A_α' we shall have the equation

$$T_\alpha' = T_1^{-\lambda_1} T_2^{-\lambda_2} \cdots T_{\alpha-1}^{-\lambda_{\alpha-1}}, \tag{3}$$

where $\lambda_1, \cdots, \lambda_{\alpha-1}$ are all divisible by p. From the remaining relations of type (2) we eliminate T_α by means of the equation $T_\alpha = T_1^{-\lambda_1\lambda_\alpha^{-1}} \cdots T_r^{-\lambda_r\lambda_\alpha^{-1}}$, arranging, of course, the elements in each new relation according to the sequence $T_1, T_2, \cdots, T_{\alpha-1}, T_{\alpha+1}, \cdots$. If none of the exponents in these new relations is prime to p, our process is at an end; otherwise, we proceed as before until we eventually determine a set of elements $A_1', \cdots, A_r'$ for which a certain h of the T''s, say $T_{e_1}', T_{e_2}', \cdots, T_{e_h}'$, constitute a set of I.G.O. for H, while each of the remaining T''s is of the form

$$T_{e_1}'^{d_1 p} \cdots T_{e_h}'^{d_h p}.$$

That $A_1', \cdots, A_r'$ constitute a U-basis for A is obvious from the fact that $|a_{ij}|$ equals $\lambda_\alpha \lambda_\alpha' \cdots$, while $\lambda_\alpha, \lambda_\alpha', \cdots$ are all prime to p. (See (a) above;

the elements below the main diagonal in $|a_{ij}|$ are all zeros.)

From Theorems I and III of §9, in connection with the equation $t_i = \theta^a s_i$, we know that for every ω-group the elements of a U-basis for G correspond to the elements of a U-basis for A, and conversely. That this correspondence is not necessarily a reciprocal one when G is not an ω-group is clear from the following example.

Let G be the metabelian group defined by the relations

$$s_1^p = s_2^p = s_0^p = E, \quad s_1^{-1}s_2s_1 = s_2s_0, \quad s_1s_0 = s_0s_1, \quad s_2s_0 = s_0s_2.$$

Let a be the smallest positive root of the congruence $2a+1\equiv 0 \pmod p$. By an easy computation we can show that $A_1 = s_1S_1^a$, $A_2 = s_2S_2^a$, $A_3 = A_1^{-1}A_2^{-1}s_0^{-a}$ constitute a U-basis for A. But A_1T_1, A_2T_2, A_3T_3 do not constitute a U-basis for G, since $A_x = (A_1T_1)^{x_1}(A_2T_2)^{x_2}(A_3T_3)^{x_3}$ is the identity for $x_1\equiv x_2\equiv x_3 \pmod p$. In fact, for $x_1\equiv x_2\equiv x_3\equiv 1 \pmod p$, A_x reduces to $s_1s_2s_1^{-1}s_2^{-1}s_0$, which is clearly the identity.

Properties associated with a given basis

11. Having demonstrated the occurrence in G of each of the four types of bases, we now propose to develop certain "non-invariant" properties which are associated with a particular choice of a basis for G. From this point on, the letters $\beta_1, \beta_2, \cdots, \beta_n$ shall represent a special kind of B-basis, namely an *MB-basis,* which we define in the following manner: With every B-basis of G there is associated a number χ, which equals the sum of the orders of the elements in this B-basis. Those B-bases for which χ is a minimum in G we shall call MB-bases.

Let the elements of any MB-basis be denoted by $\beta_1, \beta_2, \cdots, \beta_n$, of orders $p^{\eta_1}, \cdots, p^{\eta_n}$ respectively, where $\eta_1 \geqq \eta_2 \geqq \cdots \geqq \eta_n$. We know that every operation of G can be represented in the form $\beta_x' = \beta_x c$, where β_x equals $\beta_1^{x_1}\beta_2^{x_2}\cdots\beta_n^{x_n}$, while c is some element of C. Furthermore, we know that the order of β_x is the order of its constituent $\beta_i^{x_i}$ of highest order. We now prove a result which is of great importance in the following development of the theory.

THEOREM I. *If β_x is a principal element of G, then the order of β_x' is the order of that one of its constituents $\beta_1^{x_1}, \beta_2^{x_2}, \cdots, \beta_n^{x_n}$, c which is of highest order.*

The theorem is clearly true when β_x and c are of unequal orders. So we assume that β_x and c are both of order p^a, while β_x' is of order p^b, $b<a$. For the purpose of demonstrating the impossibility of the inequality $b<a$, it is permissible to assume that (a) among all the products $\beta_y' = \beta_y c$ of a principal element of G into an element of C where β_y' is of lower order than β_y, there is

none whose order is less than p^b.

Let x_α be the first one of the exponents $x_1, x_2, \cdots, x_n$ in β_x which is prime to p. We wish to show that by replacing in our given MB-basis the element β_α by β_x', we shall obtain a B-basis. Since the sum of the orders of the elements in this new basis will be less than Σp^{n_i}, $i=1, 2, \cdots, n$, we shall arrive at a contradiction, since Σp^{n_i} is a minimum in G.

Now $\beta_1, \cdots, \beta_{\alpha-1}, \beta_x', \beta_{\alpha+1}, \cdots, \beta_n$ will generate G. Hence we have only to prove that $\beta_\lambda = E$, where β_λ is $\beta_1^{\lambda_1} \cdots \beta_{\alpha-1}^{\lambda_{\alpha-1}} \beta_x'^{\lambda_\alpha} \cdots \beta_n^{\lambda_n}$, requires that each λ be divisible by the order of the element to which it belongs. If λ_α is divisible by p^b, then there is nothing to prove. So we assume that p^d, the highest power of p which divides λ_α, is less than p^b.

In β_λ we replace β_x' by $\beta_x c$ and bring the result into the form $\beta_\lambda' = \beta_1^{\lambda_1 + \lambda_\alpha x_1} \cdot \beta_\alpha^{\lambda_\alpha x_\alpha} \cdots \beta_n^{\lambda_n + \lambda_\alpha x_n} c^{\lambda_\alpha} \bar{c}$, where $\bar{c}$ is a product of commutators, each of whose exponents is divisible by λ_α. Let p^e be the highest power of p that divides every exponent in β_λ'. Clearly e is not greater than d.

Now we can find in G an operation $\beta_z' = \beta_z c'$, where β_z is a principal element $\beta_1^{z_1}\beta_2^{z_2} \cdots \beta_n^{z_n}$ in G, such that $\beta_z'^{p^e}$ equals β_λ' (see (d) of §1). The order of β_z is clearly greater than p^e; the order of β_z' is p^e, since β_λ' is the identity. This, however, involves a contradiction of assumption (a), since e is less than b. We conclude, therefore, that d must equal b. This completes the demonstration of Theorem I.

THEOREM II. *If an operation s of G can be represented in the form β_x, where each exponent x_i is a least positive residue modulo p^{n_i}, then the x_i's are uniquely determined.*

Suppose that s is given by β_x and also by β_y, where β_y is $\beta_1^{y_1}\beta_2^{y_2} \cdots \beta_n^{y_n}$. Then $\beta_x = \beta_y$ leads to $\beta_x \beta_y^{-1} = E$. This latter equation can be reduced to the form $\beta_{x-y} c_{x-y} = E$, where β_{x-y} is $\beta_1^{x_1-y_1} \cdots \beta_n^{x_n-y_n}$ and c_{x-y} is $\prod_{i<j} c_{ij}^{-y_i(x_j-y_j)}$, c_{ij} being $\beta_i^{-1}\beta_j\beta_i\beta_j^{-1}$. Our theorem will follow if we can show that β_{x-y} must be the identity, since $\beta_{x-y} = E$ requires $x_i - y_i \equiv 0 \pmod{p^{n_i}}$.

Suppose that β_{x-y} is not the identity. Then each exponent in β_{x-y} must be divisible by p; otherwise, β_{x-y} could not be in the ϕ-subgroup of G. Let p^α be the highest power of p that divides every $x_i - y_i$. Since every exponent in c_{x-y} contains one of the $x_i - y_i$, we can find in G an operation $\beta_z' = \beta_z c' = \beta_1^{z_1}\beta_2^{z_2} \cdots \beta_n^{z_n} c'$, such that β_z is a principal element of G, and such that $\lambda_z'^{p^\alpha}$ is $\beta_{x-y} c_{x-y}$. Since the order of β_z exceeds p^α, this leads to a contradiction of Theorem I. Hence β_{x-y} must be the identity.

THEOREM III. *If G is an ω-group, then every B-basis is a U-basis, and conversely.*

To prove the "conversely" we need only to show that the elements $P_1, P_2, \cdots, P_r$ of a U-basis constitute a set of I.G.O. for G. The ϕ-subgroup of an ω-group is $\mho_1(G)$ Clearly a product $P_1^{\lambda_1}P_2^{\lambda_2}\cdots P_r^{\lambda_r}$ can be a pth power in G only when each λ is divisible by p.

To prove the first part of our theorem we make use of Theorem IV of §6. Knowing that we can express every operation s of G in the form $\bar\beta_x=\bar\beta_1^{x_1}\bar\beta_2^{x_2}\cdots\bar\beta_n^{x_n}$, where $\bar\beta_1, \bar\beta_2, \cdots, \bar\beta_n$ are a B-basis for G, we have only to show that s is uniquely represented by $\bar\beta_x$, whenever the exponents x_i are least positive residues.

In the proof of Theorem II above we use the assumption that $\beta_1, \beta_2, \cdots, \beta_n$ are an MB-basis in order to show that $\beta_x=\beta_y$ requires $\beta_{x-y}=E$. But if G is an ω-group, we can prove this without requiring that the β's constitute an MB-basis. If $\bar\beta_1, \bar\beta_2, \cdots, \bar\beta_n$ are simply a B-basis, then $\bar\beta_{x-y}=E$ will hold only when each x_i-y_i is divisible by the order of the $\bar\beta_i$ to which it belongs. In the case of an ω-group every commutator can be expressed in the form $\bar\beta_1^{d_1p}\bar\beta_2^{d_2p}\cdots\bar\beta_n^{d_np}$. From this we see that $\bar\beta_{x-y}$ and $\bar c_{x-y}$ (in Theorem II) can never be of the same order unless each is the identity. That is, if G is an ω-group, then in Theorem II we may replace our assumption that $\beta_1, \beta_2, \cdots, \beta_n$ are an MB-basis by the weaker assumption that they are a B-basis.

From this modified form of Theorem II we see that the elements $\bar\beta_1, \bar\beta_2, \cdots, \bar\beta_n$ satisfy the requirements (A) and (C) of §9 and accordingly constitute a U-basis for G.

Theorem IV. *If G is an ω-group, then every B-basis is an MB-basis.*

This follows directly from Theorem III, since the orders of the elements of a U-basis are an invariant of G.

Theorem V. *The orders of the elements in any MB-basis for G are an invariant of G.*

When G is an ω-group, this follows directly from Theorem III. We let $\beta_1, \beta_2, \cdots, \beta_n$, of orders $p^{\eta_1}, p^{\eta_2}, \cdots, p^{\eta_n}$ respectively, and $\beta_1', \beta_2', \cdots, \beta_n'$, of orders $p^{\eta_1'}, p^{\eta_2'}, \cdots, p^{\eta_n'}$, be any two MB-bases for G. We may assume the inequalities $\eta_1\geqq\eta_2\geqq\cdots\geqq\eta_n$ and $\eta_1'\geqq\eta_2'\geqq\cdots\geqq\eta_n'$. Now η_1 and η_1' must be equal. Let η_α' be the first one of the η''s which differs from its corresponding η, and let η_α' be less than η_α. Now the elements $\beta_1', \cdots, \beta_n'$ can be expressed in terms of the β's by means of the equations

$$\beta_i' = \beta_1^{a_{i1}}\beta_2^{a_{i2}}\cdots\beta_n^{a_{in}}c_i \qquad (i = 1, 2, \cdots, n) \tag{1}$$

where $c_1, \cdots, c_n$ are elements of C. Since the β''s are a set of I.G.O. for G, it is obvious that the determinant $|a_{ij}|$ must be prime to p. Since the order of

β_α' is $p^{\eta_\alpha'}$, either (b) $a_{\alpha 1}, a_{\alpha 2}, \cdots, a_{\alpha\alpha}$ are divisible by $p^{\eta_1-\eta_\alpha'}, p^{\eta_2-\eta_\alpha'}, \cdots, p^{\eta_\alpha-\eta_\alpha'}$ respectively, or (c) the order of $\beta_1^{a_{\alpha 1}}\beta_2^{a_{\alpha 2}} \cdots \beta_n^{a_{\alpha n}}$ exceeds the order of β_α'. In case (b), the determinant $|a_{ij}|$ would be divisible by p, while in case (c), we should have a contradiction of Theorem I. Consequently, η_i' must equal η_i, for $i=1, 2, \cdots, n$.

To our list of invariants in §6 we may add the invariants $p^{\eta_1}, p^{\eta_2}, \cdots, p^{\eta_n}$. Obviously η_1 equals δ_1. That these invariants coincide with a certain n of the invariants $p^{\delta_1}, p^{\delta_2}, \cdots, p^{\delta_r}$ is a consequence of the following result.

THEOREM VI. *By the addition of a certain $r-n$ terms every MB-basis can be extended to a U-basis for G.*

For ω-groups the theorem is trivial. We therefore assume that G is not an ω-group.

(i) We first show that $A_1=\theta^a\beta_1$, $A_2=\theta^a\beta_2, \cdots, A_n=\theta^a\beta_n$ constitute a U-basis for the subgroup A' of A which they generate.

(ii) Next we show that we can select from A a certain $r-n$ elements $A_{n+1}, \cdots, A_r$ such that $A_1, \cdots, A_r$ will constitute a U-basis for A.

(iii) Finally, we prove that $\theta^{-a}A_1, \cdots, \theta^{-a}A_r$ constitute a U-basis for G.

Proof of (i). We have only to show that the equation

$$A_1^{\lambda_1}A_2^{\lambda_2}\cdots A_n^{\lambda_n} = E \tag{2}$$

holds only for λ_i divisible by p^{η_i}. Now $\theta^a\beta_i$ equals $\beta_iS_i^a$, where S_i transforms G according to β_i. In (2) we replace each A_i by $\beta_iS_i^a$ and bring the result into the form

$$\beta_1^{\lambda_1}\beta_2^{\lambda_2}\cdots\beta_n^{\lambda_n}\prod_{i<j} c_{ij}^{-a\lambda_i\lambda_j} = E, \tag{3}$$

where c_{ij} is $S_i^{-1}\beta_jS_i\beta_j^{-1}$.

Now $\beta_1^{\lambda_1}\beta_2^{\lambda_2}\cdots\beta_n^{\lambda_n}$ cannot be in $\Phi(G)$ unless each λ_i is divisible by p. Consequently, every exponent $a\lambda_i\lambda_j$ must be divisible by p^2. Evidently $\beta_1^{\lambda_1}\cdots\beta_n^{\lambda_n}$ must be the identity, if equation (3) is to hold. Since the β's are an MB-basis, each λ_i must be divisible by p^{η_i}. Since the order of A_i is p^{η_i}, we see that A' is the direct product of $\{A_1\}$, $\{A_2\}$, etc.

Proof of (ii). We write A in cosets with respect to A'. Let $Q_1, Q_2, \cdots, Q_{l_1}$, of orders $p^{\zeta_1}, p^{\zeta_2}, \cdots, p^{\zeta_{l_1}}$, be any U-basis for A/A'.* We wish to show that the coset of A which corresponds to $Q_j, j=1, 2, \cdots, l_1$, contains an operation of order p^{ζ_j}.

* One sees that A/A' and $C/\bar{C}_1$ have the same number of invariants. Furthermore, l_1 equals $r-n$ (see §6).

Now this coset contains an element c_j of C which is a principal element of C and is not in $\mho_1(A)$. If c_j is of order p^{ζ_j}, then we denote it by the letter A_{n+j} and add it to the set $A_1, \cdots, A_n$. If not, then there must exist an equation

$$c_j^{p^{\zeta_j}} = (A_1^{b_1} A_2^{b_2} \cdots A_n^{b_n})^{p^{\xi}}, \tag{4}$$

where the element in the parenthesis is a principal element of A'. We propose to show that ξ must exceed ζ_j. We replace, in (4), each A_i by $\beta_i S_i^a$. We may then bring (4) into the form

$$(\beta_1^{b_1} \beta_2^{b_2} \cdots \beta_n^{b_n})^{p^{\xi}} = \bar{c}^{p^{\xi}} c_j^{p^{\zeta_j}}. \tag{5}$$

(It is clear that $(S_1^{ab_1} \cdots S_n^{ab_n})^{p^{\xi}}$ must be the identity.) If ξ is not greater than ζ_j, we can determine an element c' in C such that $\beta_1^{b_1} \cdots \beta_n^{b_n} \bar{c}^{-1} c' c_j^{-p^{\zeta_j - \xi}}$ will be of order p^{ξ}. Since $\beta_1^{b_1} \beta_2^{b_2} \cdots \beta_n^{b_n}$ is a principal element of G whose order exceeds p^{ξ}, we have a contradiction of Theorem I. For the element A_{n+j} we may therefore take $c_j(A_1^{b_1} A_2^{b_2} \cdots A_n^{b_n})^{-p^{\xi - \zeta_j}}$. Obviously the r elements $A_1, \cdots, A_n, A_{n+1}, \cdots, A_{n+l_1 = r}$ constitute a U-basis for A.

Proof of (iii). Let $T_1, T_2, \cdots, T_r$ be the permutations of H which correspond by means of the equation $s_i = \theta^{-a} t_i$ to $A_1, \cdots, A_r$ as determined in (i) and (ii) above. From the manner of selection for $A_1, A_2, \cdots, A_r$ it is clear that no T_{n+j} can be a principal element of H (observe the inequality $\xi > \zeta_j$ above). Again, every product $A_{n+1}^{x_{n+1}} A_{n+2}^{x_{n+2}} \cdots A^{x_r}$ must be in $\Phi(G)$.

Now the equation

$$(A_1 T_1)^{x_1} \cdots (A_r T_r)^{x_r} = (A_1 T_1)^{y_1} \cdots (A_r T_r)^{y_r} \tag{6}$$

can be brought into the form

$$\beta_1^{x_1 - y_1} \cdots \beta_n^{x_n - y_n} = \beta_\phi, \tag{7}$$

where β_ϕ is in $\Phi(G)$. We know that (7) can exist only if each $x_i - y_i$, $i = 1, \cdots, n$, is divisible by p. We also know that (6) requires that the T's satisfy the equation

$$T_1^{x_1 - y_1} \cdots T_n^{x_n - y_n} T_{n+1}^{x_{n+1} - y_{n+1}} \cdots T_r^{x_r - y_r} = E. \tag{8}$$

Consequently, in the particular equation (8) which arises from a given equation (6) the exponent of every T which is a principal element in H must be divisible by p.* Hence the proof of Theorem IV in §9 is applicable to the U-basis $A_1, \cdots, A_r$, as determined in (i) and (ii) above. Having proved that

* In the hypothesis of Theorem IV in §9 we demanded this property of *every* equation $T_1^{x_1} \cdots T_r^{x_r} = E$. Obviously it is sufficient to require it only for that *particular* equation which arises from equation (5) of §9.

$A_1T_1, \cdots, A_rT_r$ constitute a U-basis for G, our demonstration of Theorem VI is at an end. It is, of course, obvious that the orders of the A_iT_i, viz., $p^{\eta_1}, p^{\eta_2}, \cdots, p^{\eta_n}, p^{\zeta_1}, \cdots, p^{\zeta_{t_1}}$, do not, in this sequence, necessarily coincide with $p^{\delta_1}, p^{\delta_2}, \cdots, p^{\delta_r}$ respectively.

We now mention two theorems, which are rather obvious consequences of the definition of a U-basis.

THEOREM VII. *If $P_1, P_2, \cdots, P_r$ are any U-basis for G, then the order of $P_x = P_1^{x_1}P_2^{x_2} \cdots P_r^{x_r}$ is the order of its constituent $P_i^{x_i}$ of highest order.*

THEOREM VIII. *If $P_{s_1}, P_{s_2}, \cdots, P_{s_r}$ are the elements $P_1, P_2, \cdots, P_r$ above written in any arbitrary sequence, then each element of G can be expressed uniquely in the form $P_{s_1}^{z_1}P_{s_2}^{z_2} \cdots P_{s_r}^{z_r}$, where each exponent is a least positive residue modulo the order of the element to which it belongs.*

The proofs are easily supplied.

We now prove the complement to Theorem VI.

THEOREM IX. *Let $P_1, P_2, \cdots, P_r$ be any U-basis for G. Any n elements $P_{s_1}, P_{s_2}, \cdots, P_{s_n}$ (of this U-basis) which generate G will constitute an MB-basis for G.*

Since $P_1, \cdots, P_r$ generate G, it is obvious that a certain n of them, say $P_{s_1}, \cdots, P_{s_n}$, will constitute a set of I.G.O. for G. Let the orders of these be $p^{\eta_1}, p^{\eta_2'}, \cdots, p^{\eta_n'}$, $\eta_1 \geqq \eta_2' \geqq \cdots \geqq \eta_n'$. Let $\beta_1, \beta_2, \cdots, \beta_n$, of orders $p^{\eta_1}, p^{\eta_2}, \cdots, p^{\eta_n}$, $\eta_1 \geqq \eta_2 \geqq \cdots \geqq \eta_n$, be any MB-basis for G. Our theorem will follow if we can show that η_i' must equal η_i, $i=2, 3, \cdots, n$, since $P_{s_1}, \cdots, P_{s_n}$ constitute at least a B-basis.

Now each β_i, $i=1, 2, \cdots, n$, can be expressed in the form $\beta_i = P_{s_1}^{a_{i1}}P_{s_2}^{a_{i2}} \cdots P_{s_n}^{a_{in}}c_i$, where c_i is some element of C. Suppose that η_α' is the first of the η''s, in the sequence $\eta_2', \eta_3', \cdots, \eta_n'$, which differs from its corresponding η. Since η_α' cannot be less than η_α, we take $\eta_\alpha' > \eta_\alpha$. Now $|a_{ij}|$ must be prime to p (see proof of Theorem V). Hence at least one of the exponents $a_{i1}, a_{i2}, \cdots, a_{in}$ in every β_i must be prime to p. So we take $a_{i\lambda}$ prime to p. As an element of G, c_i can be expressed uniquely in the form $P_{s_1}^{x_1}P_{s_2}^{x_2} \cdots P_{s_n}^{x_n}P_{s_{n+1}}^{x_{n+1}} \cdots$ (see Theorem VIII). In this expression each x_j, $j=1, \cdots, n$, is divisible by p, since no product $P_{s_1}^{y_1}P_{s_2}^{y_2} \cdots P_{s_n}^{y_n}$ can be in C unless each y_i is divisible by p. We can therefore express β_i in the form

$$\beta_i = P_{s_1}^{a_{i1}+b_{i1}p} \cdots P_{s_\lambda}^{a_{i\lambda}+b_{i\lambda}p} \cdots P_{s_n}^{a_{in}+b_{in}p}P_{s_{n+1}}^{b_{i,n+1}} \cdots .$$

From this we see that the order of β_i is at least equal to the order of P_{s_λ}, which is η_λ' (see Theorem VII). Consequently, for $i>j$, a_{ij} must be divisible by $p^{\eta_i'-\eta_i}$. Taking $i=\alpha$, we see that $a_{\alpha 1}, a_{\alpha 2}, \cdots, a_{\alpha\alpha}$ must be divisible by

$p^{\eta_1-\eta_\alpha}, p^{\eta_2'-\eta_\alpha}, \cdots, p^{\eta_\alpha'-\eta_\alpha}$ respectively. But for $\eta_\alpha' > \eta_\alpha$, this would lead to $|a_{ij}| \equiv 0 \pmod p$. Hence the assumption $\eta_\alpha' > \eta_\alpha$ is impossible, and the elements $P_{s_1}, P_{s_2}, \cdots, P_{s_n}$ must constitute an MB-basis for G.

Defining relations for G

12. In this section we shall develop a compact set of abstract defining relations for G which arise from the elements $P_1, P_2, \cdots, P_r$ of a given U-basis for G.

As before, we denote the orders of $P_1, P_2, \cdots$ by $p^{\delta_1}, p^{\delta_2}, \cdots$. We define the symbol $R_{p^e}(q)$ to be the least positive residue of q modulo p^e. Again, by the symbol $R[P_1^{x_1} \cdots P_r^{x_r}]$—in short, $R[P_x]$—we mean the result obtained by replacing each exponent x_i by its least positive residue modulo p^{δ_i}. That is,

$$R[P_1^{x_1} \cdots P_r^{x_r}] \tag{1}$$

is the product of r terms $P_i^{R_i}$, where $R_i = R_{p^{\delta_i}}(x_i)$. Let P_{ij} be defined by the equation $P_{ij} = P_i^{-1}P_j^{}P_iP_j^{-1}$, and let $p^{\delta_{ij}}$ denote the order of P_{ij}. We know that each P_{ij} can be represented uniquely in the form

$$P_{ij} = P_1^{b_{1ij}}P_2^{b_{2ij}} \cdots P_r^{b_{rij}}, \tag{2}$$

where the exponents are least positive residues. Although every P_{ij} is invariant in G, the constituents $P_\alpha^{b_{\alpha ij}}$ need not be separately invariant under G. We know, however, that the order of $P_k^{-1}P_\alpha^{b_{\alpha ij}}P_kP_\alpha^{-b_{\alpha ij}}$, $i, j, k = 1, 2, \cdots, r$, is less than the order of P_{ij}. From this fact we see that by using equations (2) we can ultimately bring any product P_xP_y (where the x's and y's are arbitrary integers) into the form $P_{x'} = P_1^{x_1'}P_2^{x_2'} \cdots P_r^{x_r'}$. For instance, the first step in this reduction is to bring P_xP_y into the form

$$P_1^{x_1+y_1} \cdots P_r^{x_r+y_r} \prod_{i<j} P_{ij}^{x_jy_i}.$$

Now P_x and P_y are operations of G, whether or not we regard the x's and y's as least positive residues. But if we wish to obtain a unique representation for each operation of G, we must obviously replace P_x by $R[P_x]$. In view of the inequalities $\delta_{ij} \leqq \delta_i$, $\delta_{ij} \leqq \delta_j$, it is clearly a matter of indifference, in bringing P_xP_y into the form $R[P_{x'}]$, whether we reduce exponents after each step (after adding together x_1 and y_1, for instance) or whether we make only a single reduction,—on the exponents of $P_{x'}$. Let us adopt this latter point of view with the proviso that in the course of bringing P_xP_y into the form $P_{x'}$ we drop out all elements $P_i^{\lambda_i}$ for which the exponent λ_i is formally divisible by p^{δ_i}, $i = 1, 2, \cdots, r$. This, of course, amounts to treating the x's and y's as unknowns during the process of constructing $P_{x'}$. We see,

therefore, that the exponents of $P_{x'}$ can be given in terms of the x's and y's by the equations $x_i' = x_i + y_i + f_i(x_1, \cdots, x_r, y_1, \cdots, y_r)$, $i = 1, 2, \cdots, r$, where f_i is either identically zero or a rational integral function of the x's and y's, each term of which is at least of the first degree in both x and y. In view of the congruence $x^{p^{\delta_i - 1}(p-1)} \equiv 1 \pmod{p^{\delta_i}}$, we may assume that the exponent of each x or y in f_i does not exceed $p^{\delta_i - 1}(p-1)$. Let us write P_w for $R[P_{x'}]$. Then the exponents of P_w are given by the equations

$$w_i = R_{p^{\delta_i}}(x_i + y_i + f_i) \qquad (i = 1, 2, \cdots, r).^* \tag{3}$$

Now each of the p^m operations $R[P_x]$ of G is completely characterized by the exponents

$$R_{p^{\delta_j}}(x_j) \qquad (j = 1, 2, \cdots, r).$$

Consequently, G is completely defined by the r numbers $p^{\delta_1}, p^{\delta_2}, \cdots, p^{\delta_r}$ and the equations (3) above. One readily sees that the form of the functions f_i depends, in general, upon the particular U-basis $P_1, P_2, \cdots, P_r$ which we select.

If each component x_i in the vector $v_x = (x_1, x_2, \cdots, x_r)$ is a least positive residue modulo p^{δ_i}, then v_x has p^m distinct values. Now equations (3) associate with any two vectors v_x and v_y a unique product $v_w = v_x v_y$. It is clear that under the law of multiplication defined by (3) those p^m vectors constitute a representation of G. Under the multiplication defined by

$$w_i = R_{p^{\delta_i}}(x_i + y_i)$$

they constitute a representation of A. The "divergence" of G from its conformal abelian group is measured, so to speak, by the r functions f_i.

It is worthwhile to mention two other representations of G which arise from equations (3). If in (3) we hold the y's fixed and let the x's range over all permissible values (i.e., least positive residues), then there is defined a regular permutation $\binom{v}{vv_y}$ of the p^m vectors. So we may regard (3) as defining a representation of G as a regular permutation group G_π.

If in (3) we regard the x's as unknowns and the y's as residues, then for a given set of values $y_1, \cdots, y_r$ there is defined a transformation τ_y, which is not necessarily linear. That is, (3) gives rise to a representation of G as a congruence group G_τ. It is a simple task to verify the fact that G_π and G_τ are simply isomorphic under the correspondence

$$\begin{pmatrix} v \\ vv_y \end{pmatrix} \sim \tau_y^{-1}.$$

---

* The x's and y's in equations (3) are to be regarded as unknowns; this point of view is essential for certain interpretations of (3) which we shall mention later. Of course in the computation above we are concerned only with values of the x's and y's which are least positive residues.

13. In §12 we indicated a means for constructing a set of defining relations for G, starting from a given U-basis for G. In §13 we set ourselves a similar task, with reference to the operations of a given MB-basis. First, however, we shall prove the following "existence" theorem.

THEOREM I. *Let $B_1, B_2, \cdots, B_n$ be n operations which satisfy the following conditions and no others*:

(1) *the order of B_i is p^{η_i}, $i=1, 2, \cdots, n$;*
(2) *the order of B_{ij} is $p^{\eta_{ij}}$, where B_{ij} is $B_i^{-1}B_jB_iB_j^{-1}$;*
(3) $\eta_{ij} \leqq \eta_i$, $\eta_{ij} \leqq \eta_j$;
(4) $B_iB_{jk}=B_{jk}B_i$, $i, j, k=1, 2, \cdots, n$; *the symbols* n, η_i, η_{jk} are arbitrary, but fixed, positive integers. Then $B_1, B_2, \cdots, B_n$ will generate a metabelian (or abelian) group F, whose order is $p^{\Sigma\eta_i+\eta_{jk}}$.*

It is, of course, permissible to assume $\eta_i>0$. If F exists, then the B_i plus those B_{jk} which are not the identity will surely constitute a U-basis for F. This suggests the introduction of the vector

$$v_x = (x_1, x_2, \cdots, x_n, x_{12}, x_{13}, \cdots, x_{1n}, x_{23}, \cdots, x_{2n}, \cdots, x_{n-1,n}),$$

where the x_i and the x_{jk}, $j<k$, are least positive residues modulis p^{η_i} and $p^{\eta_{jk}}$ respectively.† The symbol v_x has $n+n(n-1)/2$ components (each component for which η_{jk} is zero is represented by a zero); two symbols are to be regarded as distinct unless their components are identical. We readily see that v_x has $p^{\Sigma\eta_i+\eta_{jk}}$ distinct values. We propose to show that the symbols v_x constitute a group of this order, under the law of multiplication given by $v_w=v_xv_y$, where the components of v_w are defined by

$$(5)\quad \begin{aligned} w_i &= R_{p^{\eta_i}}(x_i + y_i) && (i = 1, 2, \cdots, n);\\ w_{jk} &= R_{p^{\eta_{jk}}}(x_{jk} + y_{jk} + x_ky_j) && (j = 1, \cdots, n;\ k = 2, \cdots, n;\ j < k). \end{aligned}$$

We outline a method for proving that the four group-postulates are satisfied. Obviously (5) associates with any two symbols v_x and v_y a unique product v_w; from (3) it is easy to show that multiplication is associative. The element v_0, for which every component is a zero, has the characteristic property of an identity: i.e., $v_0v_x=v_xv_0=v_x$. By computation, we find that the components of $(v_x)^\lambda$ are given by

$$(6)\quad x_i' = R_{p^{\eta_i}}(\lambda x_i),\ i = 1, \cdots, n;\quad x_{jk}' = R_{p^{\eta_{jk}}}\left(\lambda x_{jk} + \frac{\lambda(\lambda-1)}{2}x_jx_k\right).$$

* We justify this choice of symbols on the grounds that the B_i's will ultimately be identified with the elements of an MB-basis for a given G.

† From (2) and (4) it follows that B_{ij} must equal B_{ji}^{-1}; consequently η_{ij} equals η_{ji}. For x_{jk}, accordingly, we are justified in assuming $j<k$.

From (6) we see that $(v_x)^{n_x}$ equals v_0, where n_x is the smallest positive integer satisfying the simultaneous congruences $n_x x_i \equiv 0 \pmod{p^{\eta_i}}$; $n_x x_{jk} \equiv 0 \pmod{p^{\eta_{jk}}}$. The results of this paragraph show that the symbols v_x constitute a group.

To show that this group is metabelian (or abelian) we construct $v_s = v_x^{-1} v_y v_x v_y^{-1}$. Its components are given by

$$z_i = 0,\ i = 1, 2, \cdots, n;\quad z_{jk} = R_{p^{\eta_{jk}}}(x_j y_k - x_k y_j). \tag{7}$$

By referring to (5) we readily see that the commutator v_s is commutative with every v_x.

It remains to associate the symbols B_i and B_{jk} with the symbols v_x. We define v_i, $i = 1, 2, \cdots, n$, to be that vector for which the component x_i is 1 while the remaining components are zeros. We define v_{jk} as that vector for which the component x_{jk} is 1 while the remaining components are zeros. From (6) it follows that the order of each v_i is p^{η_i}, while the order of each v_{jk} is $p^{\eta_{jk}}$. From (7) we observe that v_{jk} and $v_j^{-1} v_k v_j v_k^{-1}$ are the same. As symbols, therefore, v_i and B_i are interchangeable; the same is true of v_{jk} and B_{jk}. This completes the proof of Theorem I.

Let us now assume that the numbers n, η_i, η_{jk} are no longer arbitrary, but represent respectively the number of I.G.O., the order of β_i, the order of $c_{jk}(=\beta_j^{-1}\beta_k\beta_j\beta_k^{-1})$, where $\beta_1, \beta_2, \cdots, \beta_n$ are the elements of a given MB-basis for a given metabelian group G. We construct the group F, as in Theorem I above. Each of its operations is given uniquely by the symbol

$$B_x = B_1^{x_1} B_2^{x_2} \cdots B_n^{x_n} \prod_{i<j} B_{ij}^{x_{ij}},$$

where the exponents are least positive residues.* Let ψ be defined as the operation of replacing in B_x each B_i by β_i and each B_{jk} by c_{jk}. That is, $\psi(B_x) = \beta_x$, where β_x is

$$\beta_1^{x_1} \cdots \beta_n^{x_n} \prod_{j<k} c_{ij}^{x_{ij}}.$$

We know that every operation of G is representable (although not necessarily uniquely) in the form β_x.

We state without proof two results, whose verification presents no difficulty: (a) the number of formally distinct representations of a given element σ of G in the form β_x equals the number of formally distinct representations of the identity of G; (b) the operation ψ defines an isomorphism of F with G. Let F_1 denote that subgroup of F which corresponds to the identity of G in

* We agree always to write the factors of $\Pi B_{ij}^{x_{ij}}$ in the same order, although the particular order which we adopt is clearly a matter of indifference; we furthermore agree that those B_{jk} for which η_{jk} is zero shall not occur in $\Pi B_{ij}^{x_{ij}}$.

this isomorphism. Of course, G is simply isomorphic with F/F_1.

Let $N_1, N_2, \cdots, N_{n_1}$ denote an MB-basis for F_1. Of course, each N_i is representable uniquely in the form

$$N_i = B_1^{\xi_{i1}} \cdots B_n^{\xi_{in}} \prod_{j<k} B_{jk}^{\xi_{ijk}} \qquad (i = 1, \cdots, n_1). \tag{8}$$

By applying the operation ψ to every element of F_1 we obviously get all the formally distinct representations $\beta_x(E)$ of the identity of G. We readily see that every relation $\beta_x = E$ in G can be derived from the n_1 relations

$$\beta_1^{\xi_{i1}} \cdots \beta_n^{\xi_{in}} \prod_{j<k} c_{jk}^{\xi_{ijk}} = E \qquad (i = 1, \cdots, n_1).* \tag{9}$$

Now the data given by (1), $\cdots$, (4), (9) were determined from the elements of a particular MB-basis for G. With these data at hand we are able to identify any product of powers of the *symbols* β_i and c_{jk} with a definite element of G. In this sense, therefore, we may regard (1), $\cdots$, (4) and (9) as constituting a set of abstract defining relations for G.

14. In this section we shall outline a third method for defining G, which amounts to a specialization of the procedure developed in §§3–4. The notation which we shall employ is the same as that given in §4 and in §9, assuming, of course, that G is represented as a regular permutation group.

Let $A_1, \cdots, A_r$, of orders $p^{\delta_1}, \cdots, p^{\delta_r}$, be the elements of any U-basis for A, and let A_{jk} be defined by the equation $A_{jk} = T_j^{-1} A_k T_j A_k^{-1}$, $j, k = 1, 2, \cdots, r$. As an element in A, A_{jk} may be represented in the form

$$A_{jk} = A_1^{a^j_{k1}} A_2^{a^j_{k2}} \cdots A_r^{a^j_{kr}} \tag{1}$$

(the letter j in a^j_{k1} is a superscript, not an exponent). Since each operation T_j replaces A_k by $A_k A_{jk}$, we may represent T_j by the substitution

$$\begin{aligned} A_1 \rightarrow A_1' &= A_1^{a^j_{11}+1} A_2^{a^j_{12}} \cdots A_r^{a^j_{1r}}, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ A_r \rightarrow A_r' &= A_1^{a^j_{r1}} A_2^{a^j_{r2}} \cdots A_r^{a^j_{rr}+1}. \end{aligned} \tag{2}$$

The matrix of the exponents in (2) we shall call M_j.

We know that the operation θ^{-a} (see §3) when applied in turn to each element of A yields all the operations of G. In particular,

$$\theta^{-a} A_1^{x_1} A_2^{x_2} \cdots A_r^{x_r} \text{ is } (A_1^{x_1} \cdots A_r^{x_r})(T_1^{x_1} \cdots T_r^{x_r}).$$

---

* In connection with the various representations $\beta_x(E)$ of the identity of G, which we may denote by $\beta_y' c_y''$, where

$$\beta_y' = \beta_1^{y_1} \cdots \beta_n^{y_n} \text{ and } c_y'' = \Pi c_{jk}^{y_{jk}},$$

it is of some interest to note that the order of β_y' cannot exceed the order of c_y''. If $\beta_y' c_y''$ is one of the n_1 elements $\psi(N_i)$, $i = 1, 2, \cdots, n_1$, then either β_y' is the identity or β_y' and c_y'' are of the same order.

Since each T_j is completely characterized by its matrix M_j, it follows that G is defined by the orders of $A_1, A_2, \cdots, A_r$, the exponents in the r matrices $M_1, M_2, \cdots, M_r$, and the operation θ^{-a}.*

From the known properties of G we may state certain necessary conditions which the elements of M_j must fulfill. Since

(3′) the order of T_j divides the order of A_j,

(3) each $\overset{j}{a}_{ik}$ is divisible by $p^{\delta_k-\delta_i}$ for $i>k$;

since

(4′) every A_{jk} is commutative with every T_i,

(4) $\sum \overset{u}{a}_{jk}\overset{v}{a}_{kl}$ must be a multiple of p^{δ_l} where j, k, l, u, v range independently from 1 to r. From the equality

(5′) $A_{jk}=A_{kj}^{-1}$

we obtain

(5) $\overset{v}{a}_{ul}+\overset{u}{a}_{vl}\equiv 0 \pmod{p^{\delta_l}}$, $u, v, l=1, 2, \cdots, r$.

As a special case of (5) we have

(6) $\overset{i}{a}_{ij}\equiv 0 \pmod{p^{\delta_j}}$, $i, j=1, \cdots, r$,

which may be derived immediately from the fact that

(6′) T_i is commutative with A_i, $i=1, 2, \cdots, r$.

From the conclusions of the paragraph above we derive two additional results:

(7) the matrix of the exponents in T_j^x is given by

$$(M_j)^x = \begin{pmatrix} \overset{i}{a}_{11}x+1 & \overset{i}{a}_{12}x \cdots \overset{i}{a}_{1r}x \\ \cdots & \cdots \\ \overset{i}{a}_{r1}x & \overset{i}{a}_{r2}x \cdots \overset{i}{a}_{rr}x+1 \end{pmatrix};$$

(8) the matrix for $T_1^{x_1}T_2^{x_2}\cdots T_r^{x_r}$ is given by

$$\begin{pmatrix} \sum_{k=1}^{r}\overset{k}{a}_{11}x_k+1 \cdots \sum_{k=1}^{r}\overset{k}{a}_{1r}x_k \\ \cdots\cdots\cdots \\ \sum_{k=1}^{r}\overset{k}{a}_{r1}x_k \quad \cdots \sum_{k=1}^{r}\overset{k}{a}_{rr}x_k+1 \end{pmatrix}.$$

The foregoing results, as well as the symbols involved, are based on the assumption that we are given a regular permutation group G. The operations θ and T_j, as originally defined, have a meaning only when every permutation

* For this method of defining G it is clearly a matter of indifference whether or not the elements

$$\theta^{-a}A_i=A_iT_i \qquad (i=1, 2, \cdots, r)$$

constitute a U-basis for G.

of G is regarded as known. We wish to reinterpret the operations T_j quite apart from the assumed existence of G, under the sole assumption that $A_1, \cdots, A_r$ are a U-basis for a given abstract abelian group A. (We do not think of A as having any particular concrete representation.) As above, we shall denote the orders of $A_1, \cdots, A_r$ by $p^{\delta_1}, \cdots, p^{\delta_r}$ respectively.

We now define $T_j, j=1, \cdots, r$, to be the substitution $A_1 \to A_1', \cdots, A_r \to A_r'$, which is given by (2) above. For this substitution to define an automorphism of A, it is necessary and sufficient that the r^2 elements a_{ik}^j be integers which satisfy the following two conditions: (a) the determinant $|M_j|$ of M_j is prime to p; (b) for $i>k$, a_{ik}^j is divisible by $p^{\delta_k-\delta_i}$. Let us assume that the elements of M_j have any integral values which satisfy (3), (4), and (5) above. Since (3) and (b) are identical, in order to show that T_j now defines an automorphism of A, it is sufficient to prove that $|M_j|$ is prime to p. This we can derive as a consequence of (4). Or, from (7), which was derived from (4), we see that some power of M_j is the identity matrix, whence $|M_j|$ must surely be prime to p. As a consequence of these restrictions which we have imposed on the elements of M_j, it follows that the operations $T_1, \cdots, T_r$ may be interpreted as automorphisms of A.

In the course of verifying that (3) follows from (3′), (4) from (4′), (5) from (5′) it becomes evident that these three statements are reversible, in the sense that (3′), (4′), (5′) as a whole follow from (3), (4), (5). Therefore, the r automorphisms $T_1, T_2, \cdots, T_r$ generate an abelian group, and A is isomorphic with this abelian group under the correspondence defined by

$$A_i \sim T_i \qquad (i = 1, 2, \cdots, r). \tag{9}$$

By applying the theorem in the second footnote to §4, we conclude that the products A_xT_x, where

$$A_x \text{ is } A_1^{x_1}A_2^{x_2} \cdots A_r^{x_r} \text{ and } T_x \text{ is } T_1^{x_1}T_2^{x_2} \cdots T_r^{x_r},$$

constitute a group $\overline{G}$ of order $p^{\Sigma\delta_i}$. That this group is metabelian follows from (4′) and the fact that the commutator subgroup of $\overline{G}$ is generated by the A_{jk}. That G is conformal with A follows from (3′), (6′), and (8).

We append a rough summary of this section. In the first part we showed that for a given G and a given U-basis for A there is determined a set of elements for each of the r matrices M_j, the elements being uniquely determined if we require that each a_{ik}^j be a least positive residue modulo p^{δ_k}. These matrices, together with the orders of $A_1, \cdots, A_r$, define G, since each element of G can be given in the form A_xT_x. We enumerated certain necessary conditions which the elements of these matrices must satisfy. In the second part of this section we proved that these "necessary conditions" are

"sufficient," in the sense that any choice of elements for M_j which is consistent with these conditions will give rise to an automorphism T_j of A, while the totality of distinct products $A_x T_x$ will constitute a metabelian group conformal with A. This latter part of §14 represents a refinement of the procedure given in §4. The problem of constructing all metabelian groups which are conformal with a given abelian group A is therefore reduced to the problem of determining all possible sets of elements a_{ik}^{j} which satisfy conditions (3), (4), and (5) above.*

15. In this section we shall outline a method for constructing a representation of G as a linear congruence group. No proofs will be given, since the demonstrations in every instance follow obvious and familiar lines. We assume that G is given as a regular permutation group, and we assume that $A_1, \cdots, A_r$ are any primary U-basis for A. Then the elements in the r matrices M_j (see §14) are determined; with no loss of generality, we may assume that each a_{ik}^{j} is a least positive residue modulo p^{δ_k}.

Now the permutation A_1 may be represented by the symbol

$$\begin{pmatrix} A_1^{x_1} & A_2^{x_2} \cdots A_r^{x_r} \\ A_1^{x_1+1} & A_2^{x_2} \cdots A_r^{x_r} \end{pmatrix}.$$

Since T_1 is commutative with A_1, we may designate T_1 by the symbol

$$\begin{pmatrix} A_1^{x_1} & A_2^{x_2} \cdots A_r^{x_r} \\ A_1^{x_1} & A_2'^{x_2} \cdots A_r'^{x_r} \end{pmatrix},$$

where $A_2', \cdots, A_r'$ are defined by (2) of §14. The product $A_1 T_1$ is given by

$$\begin{pmatrix} A_1^{x_1} & A_2^{x_2} \cdots A_r^{x_r} \\ A_1^{x_1+1} & A_2'^{x_2} \cdots A_r'^{x_r} \end{pmatrix}. \tag{1}$$

In (1) we replace $A_2', \cdots, A_r'$ by their equivalents from (2) of §14 and bring the result into the form

$$\begin{pmatrix} A_1^{x_1} & A_2^{x_2} \cdots A_r^{x_r} \\ A_1^{x_1'} & A_2^{x_2'} \cdots A_r^{x_r'} \end{pmatrix}. \tag{2}$$

Now the permutation $A_1 T_1$ is completely characterized by the exponents in (2) above. Hence $A_1 T_1$ is defined by the linear substitution

$$Z_1: \begin{array}{lll} z_1' \equiv z_1 + a_{21}^{1} z_2 + a_{31}^{1} z_3 + \cdots + a_{r1}^{1} z_r + 1 & & \pmod{p^{\delta_1}}, \\ z_2' \equiv (a_{22}^{1} + 1) z_2 + a_{32}^{1} z_3 + \cdots + a_{r2}^{1} z_r & & \pmod{p^{\delta_2}}, \\ \cdots\cdots\cdots\cdots\cdots\cdots & & \\ z_r' \equiv a_{2r}^{1} z_2 + a_{3r}^{1} z_3 + \cdots + (a_{rr}^{1} + 1) z_r & & \pmod{p^{\delta_r}}. \end{array} \tag{3}$$

* We naturally restrict our choice of the a_{ik}^{j} to those which are least positive residues modulo p^{δ_k}.

By an analogous procedure we may construct the linear substitutions $Z_2, \cdots, Z_r$ which define $A_2T_2, \cdots, A_rT_r$ respectively. Since $A_1T_1, \cdots, A_rT_r$ generate G (see Theorem IV of §9), $Z_1, Z_2, \cdots, Z_r$ will generate a representation of G as a linear congruence group.

Brown University,
Providence, R.I.

A PROBLEM CONCERNING ORTHOGONAL POLYNOMIALS*

BY

G. SZEGÖ

INTRODUCTION

In his paper *Note on the orthogonality of Tchebycheff polynomials on confocal ellipses*,† Walsh has obtained a new orthogonality property of the Tchebycheff polynomials cos k arc cos z arising by orthogonalization of the set $1, z, z^2, \cdots$ over the range $-1 \leqq z \leqq 1$ with the weight function $|1-z^2|^{-1/2}$. Walsh showed that these polynomials have the same orthogonality property on all confocal ellipses with the foci at ± 1 and with the same weight function $|1-z^2|^{-1/2}$. Another example of this kind is the set of concentric circles with the weight function 1: the corresponding orthogonal polynomials $1, z, z^2, \cdots$ are the same for all curves of this set, provided the common center of the circles is at the origin.

Walsh raised the question whether there exist other pairs of curves with suitable weight functions such that the corresponding orthogonal polynomials would differ only by constant factors. A complete answer to this question seems to be rather intricate. The following theorem may furnish some indications as to the possibilities to be expected.

THEOREM 1. *Let C_1 and C_2 be two analytic Jordan curves, $n_1(z)$ and $n_2(z)$ any corresponding weight functions, positive and continuous, and*

$$p_0(z),\ p_1(z),\ p_2(z),\ \cdots,\ p_k(z),\ \cdots$$

a system of polynomials, the exact degree of $p_k(z)$ being k, simultaneously orthogonal on either of the curves,

$$\int_{C_1} n_1(z)\ p_k(z)\overline{p_l(z)}\,|dz| = \int_{C_2} n_2(z)p_k(z)\overline{p_l(z)}\,|dz| = 0, \qquad k \neq l.$$

Then one of the curves, say C_1, must contain the other (C_2) and C_1 is a level curve in the conformal mapping of the region outside C_2 onto the exterior of a circle, the points at infinity corresponding to each other. Further there is an analytic function $D(z)$ regular and non-vanishing outside C_2, $z=\infty$ inclusive, such that

* Presented to the Society, December 29, 1934; received by the editors June 29, 1934.

† Bulletin of the American Mathematical Society, vol. 40 (1934), pp. 84–88.

$$|D(z)|^2 = n_1(z), \; z \text{ on } C_1; \quad \lim_{z\to z_0} |D(z)|^2 = n_2(z_0), \; z_0 \text{ on } C_2;$$

in the second formula z_0 is an arbitrary point on C_2 and z tends to z_0 remaining in the region outside C_2.

This theorem is valid also under more general assumptions. For the sake of simplicity we confine ourselves to analytic curves. The proof is a slight extension of a known line of argument used in several papers of the author.*

The result stated above suggests in a natural way the following

PROBLEM. *To determine all Jordan curves C and all analytic functions $D(z)$ regular and non-vanishing outside C, $z=\infty$ inclusive, possessing the following property. Let C_r be a level curve in the conformal mapping of the region exterior to C onto the region exterior to the circle $|w|=r_0$, the points at infinity corresponding to each other. The orthogonal polynomials*

$$p_0(z),\, p_1(z),\, \cdots,\, p_k(z),\, \cdots$$

associated with C_r and with the weight function $|D(z)|^2$ are independent of r for $r>r_0$. In other words it is required that

$$\int_{C_r} |D(z)|^2 p_k(z)\overline{p_l(z)}\, |dz| = 0, \qquad k \neq l, \qquad r > r_0.$$

This problem admits of a complete solution. The present paper is devoted to the enumeration of all the types satisfying the condition stated above. There are altogether five essentially distinct cases, two of which have been already mentioned above. In all these cases a linear transformation of the variable z and a multiplication of the weight function by a positive constant factor of course are still allowed. The orthogonal polynomials are not necessarily normalized, the normalizing factor being in general different for different curves C_r. The five types in question are as follows.

I. C_r is the set of concentric circles $|z|=r$, $r>0$;

$$D(z) = 1, \quad p_k(z) = z^k.$$

II. C_r is the set of concentric circles $|z|=r$, $r>1$;

$$D(z) = (1 - z^{-\alpha})^{-1}, \; \alpha \text{ a positive integer},$$

$$p_k(z) = z^k, \quad 0 \leqq k < \alpha; \quad p_k(z) = z^{k-\alpha}(z^\alpha - 1), \quad k \geqq \alpha.$$

* *Beiträge zur Theorie der Toeplitzschen Formen*, II, Mathematische Zeitschrift, vol. 9 (1921), pp. 167–190, especially p. 178; *Über orthogonale Polynome, die zu einer gegebenen Kurve der komplexen Ebene gehören*, Ibidem, vol. 9 (1921), pp. 218–270, especially pp. 260–262.

III. C_r is the set of confocal ellipses with foci at ± 1;

$$D(z) = \{z + (z^2 - 1)^{1/2}\}^{1/2}(z^2 - 1)^{-1/4} = \{\tfrac{1}{2}(1 - w^{-2})\}^{-1/2},$$

$$p_k(z) = w^k + w^{-k}, \ 2z = w + w^{-1}, \ |w| = r > 1.$$

IV. C_r is the same set as in III;

$$D(z) = \{z + (z^2 - 1)^{1/2}\}^{-1/2}(z^2 - 1)^{1/4} = \{\tfrac{1}{2}(1 - w^{-2})\}^{1/2},$$

$$p_k(z) = (w^{k+1} - w^{-k-1})/(w - w^{-1}),$$

where w is the same as in III.

V. C_r is the same as in III;

$$D(z) = (z - 1)^{1/4}(z + 1)^{-1/4} = (1 - w^{-1})^{1/2}(1 + w^{-1})^{-1/2},$$

$$p_k(z) = (w^{k+1/2} - w^{-k-1/2})/(w^{1/2} - w^{-1/2}).$$

It should be observed that Tchebycheff polynomials III, in addition to the property discussed here, have another analogous one, viz. that they minimize the $\max|z^k+a_1z^{k-1}+ \cdots|$ on all the ellipses defined above. This property which was pointed out by Faber,* is analogous to that obtained by Walsh. Our line of argument given in §I applies without difficulty to Tchebycheff polynomials minimizing the max $n(z)|z^k+a_1z^{k-1}+ \cdots|$ on prescribed curves, $n(z)$ being a given weight function, positive and continuous; thus for this problem we are lead to a theorem analogous to Theorem 1.

In §I we prove Theorem 1 concerning the question raised by Walsh. §II contains a short discussion of the polynomials enumerated under I–V, particularly with respect to their orthogonality. In §III we deal with the principal problem and prove that the only possible polynomials orthogonal on all level curves of a conformal mapping are those of §II.†

I. Proof of Theorem 1

1. Let us consider an analytic Jordan curve C with a positive and continuous weight function $n(z)$. There is no difficulty in showing the existence of an analytic function $D(z)$ regular and non-vanishing outside C, $z=\infty$ inclusive, with the boundary property

$$\lim_{z\to z_0} |D(z)|^2 = n(z_0),$$

* G. Faber, *Über Tschebyscheffsche Polynome*, Journal für Mathematik, vol. 150 (1920), pp. 79–106, especially pp. 84–86.

† After having completed this paper, I communicated its main results to Professor Walsh who kindly informed me that he also obtained the first part of Theorem 1 and proposed precisely the same problem as stated above, without discussing it. These results of Walsh will appear in a monograph of the Mémorial series, Paris, Gauthier-Villars, under the title *Approximation by Polynomials in the Complex Domain*. Nevertheless, for the sake of completeness, we give here a short proof of Theorem 1.

where z_0 denotes an arbitrary point of C and z tends to z_0 from the exterior of C. The function $D(z)$ is completely determined up to a factor of the absolute value 1.

The proof of this statement can be based on the conformal mapping of the region outside C onto the region outside the unit circle $|w|=1$, the points at infinity corresponding to each other. The mapping function and its inverse,

$$z = g(w) = gw + g_0 + g_1 w^{-1} + g_2 w^{-2} + \cdots,$$
$$w = \gamma(z) = \gamma z + \gamma_0 + \gamma_1 z^{-1} + \gamma_2 z^{-2} + \cdots, \quad g\gamma = 1,$$

are uniquely determined under the assumption $g>0$. We write

$$D[g(w)] = \Delta(w), \quad |w| > 1; \; n[g(w)] = \nu(w), \quad |w| = 1.$$

Then the function $\Delta(w)$ must satisfy the following condition:

$$\lim_{w\to w_0} |\Delta(w)|^2 = \nu(w_0); \quad |w_0| = 1, \quad |w| > 1,$$

from which it can be computed by means of the Poisson integral

$$2 \log \Delta(w) = (2\pi)^{-1} \int_{-\pi}^{\pi} \log \nu(e^{i\phi}) \frac{w + e^{i\phi}}{w - e^{i\phi}} d\phi.$$

2. Let $p_k(z) = p_k z^k + \cdots$ denote the orthogonal polynomials associated with C and with the weight function $n(z)$, the normalization being arbitrary. Then it is known* that the minimum μ_k^2 of the integral

$$(2\pi)^{-1} \int_C n(z) \, | z^k + a_1 z^{k-1} + \cdots + a_k |^2 \, |dz|$$

over the set of polynomials of degree k and with the highest coefficient 1 is attained for the polynomial $p_k^{-1} p_k(z)$.

We show first that

$$\lim_{k\to\infty} \mu_k g^{-k-1/2} = |D(\infty)|. \tag{1}$$

Indeed we have

$$\begin{aligned} \mu_k^2 &= \min (2\pi)^{-1} \int_{|w|=1} |\Delta(w)|^2 \, | g(w)^k + a_1 g(w)^{k-1} + \cdots + a_k |^2 \, |g'(w)| \, |dw| \\ &= \min (2\pi)^{-1} \int_{|w|=1} |\Delta(w) \{ (g(w)/w)^k + a_1 w^{-1} (g(w)/w)^{k-1} \\ &\qquad + \cdots + a_k w^{-k} \} g'(w)^{1/2} |^2 \, |dw|, \end{aligned}$$

* See for example the second paper of the author quoted above, p. 231.

where the integrals should be interpreted as the limits of the corresponding integrals over the circle $|w|=r$ as $r\to 1+0$. The function under the absolute value sign is regular for $|w|>1$. Hence, we get a lower estimate for μ_k^2,

$$\mu_k^2 \geqq |\Delta(\infty)g^k g^{1/2}|^2 = |D(\infty)|^2 g^{2k+1}.$$

An upper estimate for μ_k^2 can be obtained for instance by using the polynomials $f_k(z)=\gamma^k z^k+\cdots$ introduced by Faber* as the principal parts of the expansions of $\gamma(z)^k$, $k=0, 1, 2, \cdots$. Faber shows by elementary methods that

$$\lim_{k\to\infty} f_k(z)\gamma(z)^{-k} = 1$$

is valid uniformly outside a level curve $|\gamma(z)|=\rho$, $\rho<1$, provided ρ is sufficiently near to 1. Now, as a consequence of the minimal property we have

$$\mu_k^2 \leqq (2\pi)^{-1}\int_C n(z)\,|q_k(z)|^2\,|dz|,$$

where $q_k(z)$ is an arbitrary polynomial of degree k with the highest coefficient 1. We put

$$q_k(z) = \sum_{h=0}^{m} \alpha_h\gamma^{h-k}f_{k-h}(z), \quad k \geqq m, \quad \alpha_0 = 1,$$

where m and the constants α_h are to be specified later. Using the asymptotic estimate above of the $f_k(z)$ we obtain

$$\begin{aligned}\overline{\lim_{k\to\infty}}\,(\mu_k g^{-k-1/2})^2 &\leqq \overline{\lim_{k\to\infty}}\,(2\pi)^{-1}\int_C n(z)\left|\sum_{h=0}^{m}\alpha_h\gamma^{h-k}g^{-k-1/2}\gamma(z)^{k-h}\right|^2|dz|\\ &\leqq (2\pi)^{-1}\int_C n(z)\left|\sum_{h=0}^{m}\alpha_h g^{-h-1/2}\gamma(z)^{-h}\right|^2|dz|\\ &\leqq (2\pi)^{-1}\int_{|w|=1}\left|\Delta(w)\sum_{h=0}^{m}\alpha_h g^{-h-1/2}w^{-h}g'(w)^{1/2}\right|^2|dw|.\end{aligned}$$

We now choose for the polynomial

$$\sum_{h=0}^{m}\alpha_h g^{-h}w^{-h} = 1+\cdots$$

the mth partial sum of the power series expansion of the analytic function

$$(g/g'(w))^{1/2}(\Delta(\infty)/\Delta(w)).$$

By taking m sufficiently large the deviation of the last integral above from

* Loc. cit., p. 83.

$|\Delta(\infty)|^2$ can be made arbitrarily small (Schwarz's inequality). Thus (1) is established.

3. The following formula is merely another expression of (1):

$$\lim_{k\to\infty}(2\pi)^{-1}g^{-2k-1}\int_C n(z)\,|\,p_k^{-1}p_k(z)\,|^2\,|\,dz\,|$$

$$= \lim_{k\to\infty}(2\pi)^{-1}g^{-2k-1}\int_{|w|=1}|\,\Delta(w)p_k^{-1}p_k[g(w)]g'(w)^{1/2}\,|^2\,|\,dw\,| = |\,\Delta(\infty)\,|^2.$$

On putting

$$g^{-k-1/2}\Delta(w)p_k^{-1}p_k[g(w)]g'(w)^{1/2} = \Delta(\infty)w^k + a_1^{(k)}w^{k-1} + a_2^{(k)}w^{k-2} + \cdots$$

we may write this as

$$\lim_{k\to\infty}(\,|\,a_1^{(k)}\,|^2 + |\,a_2^{(k)}\,|^2 + \cdots) = 0.$$

Hence we get, uniformly for $|w| \geqq r, r>1$,

$$\lim_{k\to\infty}(a_1^{(k)}w^{-1} + a_2^{(k)}w^{-2} + \cdots) = 0.$$

This yields the asymptotic formula

$$p_k(z) \sim p_k g^{k+1/2}D(\infty)\gamma(z)^k\gamma'(z)^{1/2}D(z)^{-1} \tag{2}$$

which is valid uniformly outside an arbitrary level curve C_r, $r>1$.

This formula shows immediately that the set $p_k(z)$ uniquely determines the mapping function $\gamma(z)$ as well as the function $D(z)$. The proof of Theorem 1 is thus complete.

II. Five types of orthogonal polynomials

1. It is well known that on an arbitrary circle $|z| = r$,

$$\int_{|z|=r} z^k\bar{z}^l\,|\,dz\,| = 0, \qquad k \neq l. \tag{3}$$

This equation is valid for arbitrary integral values of k and l.

2. The polynomials listed under II, in the special case $\alpha=1$, were introduced by the author.* Their orthogonality may be verified in the following manner. On putting $|z| = r>1$ we have

$$\int_{|z|=r} z^{k-1}(z-1)\bar{z}^{l-1}(\bar{z}-1)\,|\,1-z^{-1}\,|^{-2}\,|\,dz\,| = r^2\int_{|z|=r} z^{k-1}\bar{z}^{l-1}\,|\,dz\,| = 0$$

$$(k \geqq 1,\ l \geqq 1,\ k \neq l).$$

* *Über trigonometrische und harmonische Polynome*, Mathematische Annalen, vol. 79 (1919), pp. 323–339, especially p. 324.

Further for $k \geqq 1$, $z = re^{i\phi}$, $r > 1$,

$$\int_{|z|=r} z^{k-1}(z-1)\,|1-z^{-1}|^{-2}\,|dz| = r^3 \int_{-\pi}^{\pi} \frac{z^{k-1}}{\bar{z}-1}\,d\phi = r^3 \int_{-\pi}^{\pi} \frac{z^k}{r^2 - z}\,d\phi = 0.$$

In the case $\alpha > 1$ the vanishing of the integral for $k \geqq \alpha$, $l \geqq \alpha$, $k \neq l$ can be shown in the same way, and is trivial for $k < \alpha$, $l < \alpha$, $k \neq l$. The only fact which still remains to be proved is that for $k \geqq \alpha$, $l < \alpha$,

$$\int_{|z|=r} z^{k-\alpha}(z^\alpha - 1)\bar{z}^l\,|1 - z^{-\alpha}|^{-2}\,|dz| = r^{2\alpha+1} \int_{-\pi}^{\pi} \frac{z^{k-\alpha}\bar{z}^l}{\bar{z}^\alpha - 1}\,d\phi$$

$$= r^{2\alpha+1} \int_{-\pi}^{\pi} \frac{z^k}{r^{2\alpha} - z^\alpha}\left(\frac{r^2}{z}\right)^l d\phi = 0$$

which is easily verified.

3. Type III has been treated by Walsh. The proof can be presented in the following simple way. We have

$$|dz| = |\tfrac{1}{2}(1 - w^{-2})|\,|dw|,$$

and for $r > 1$, $k \neq l$, in view of (3),

$$\int_{|w|=r} (w^k + w^{-k})(\bar{w}^l + \bar{w}^{-l})\,|dw| = 0.$$

In case IV we have only to show that for $k \neq l$

$$\int_{|w|=r} \frac{w^{k+1} - w^{-k-1}}{w - w^{-1}}\,\frac{\bar{w}^{l+1} - \bar{w}^{-l-1}}{\bar{w} - \bar{w}^{-1}}\left|\frac{1 - w^{-2}}{2}\right|^2 |dw|$$

$$= \frac{1}{4} r^{-2} \int_{|w|=r} (w^{k+1} - w^{-k-1})(\bar{w}^{l+1} - \bar{w}^{-l-1})\,|dw| = 0,$$

which again follows from (3).

Finally, in case V, for $k \neq l$,

$$\frac{1}{2}\int_{|w|=r} \frac{w^{k+1/2} - w^{-k-1/2}}{w^{1/2} - w^{-1/2}}\,\frac{\bar{w}^{l+1/2} - \bar{w}^{-l-1/2}}{\bar{w}^{1/2} - \bar{w}^{-1/2}}\,|1 - w^{-1}|^2\,|dw|$$

$$= \frac{1}{2} r^{-1} \int_{|w|=r} (w^{k+1/2} - w^{-k-1/2})(\bar{w}^{l+1/2} - \bar{w}^{-l-1/2})\,|dw| = 0$$

if we use the equation

$$\int_{|w|=r} w^{k+1/2}\bar{w}^{l+1/2}\,|dw| = 0$$

which is valid for arbitrary integral values of k and l, provided $k \neq l$.

III. Solution of the principal problem

1. With the notation of §I our assumption can be written in the form

$$\int_{|w|=r} p_k[g(w)]\overline{p_l[g(w)]}\,|\Delta(w)g'(w)^{1/2}|^2\,|dw| = 0, \quad k \neq l, \quad r > 1.$$

As the first step of the proof we shall derive a power series expansion of the form

$$\begin{aligned}\Delta(w)g'(w)^{1/2}p_k[g(w)] &= \lambda_k w^k + \lambda_{k1}w^{-1} + \lambda_{k2}w^{-2} + \cdots \\ &= \lambda_k w^k + Q_k(w),\end{aligned} \tag{4}$$

where $\lambda_k \neq 0$. Let Aw^m be the second term of the power series expansion of the left-hand member of this formula. By hypothesis, if $0 \leqq m < k$, we must have

$$\int_{|w|=r} (\lambda_k w^k + Aw^m + \cdots)(\bar{\lambda}_m \bar{w}^m + \cdots)\,|dw| = 0, \quad r > 1.$$

Here the principal term for large values of r is obviously

$$A\bar{\lambda}_m \int_{|w|=r} w^m \bar{w}^m\,|dw| = 2\pi A\bar{\lambda}_m r^{2m+1}.$$

Hence $A = 0$ and the desired result follows. By (3) and the orthogonality condition we have for $k \neq l$,

$$\int_{|w|=r} (\lambda_k w^k + Q_k(w))(\bar{\lambda}_l \bar{w}^l + \overline{Q_l(w)})\,|dw|$$

$$= \int_{|w|=r} \lambda_k \bar{\lambda}_l w^k \bar{w}^l\,|dw| + \int_{|w|=r} Q_k(w)\overline{Q_l(w)}\,|dw| = 0.$$

Hence

$$\int_{|w|=r} Q_k(w)\overline{Q_l(w)}\,|dw| = 0, \quad k \neq l,$$

or

$$\lambda_{k1}\bar{\lambda}_{l1}r^{-1} + \lambda_{k2}\bar{\lambda}_{l2}r^{-3} + \cdots = 0, \quad k \neq l.$$

Consequently

$$\lambda_{k1}\bar{\lambda}_{l1} = \lambda_{k2}\bar{\lambda}_{l2} = \cdots = 0, \quad k \neq l.$$

Thus we see that no column λ_{kh} $(k=0, 1, 2, \cdots)$ of the matrix (λ_{kh}) $(k=0, 1, 2, \cdots; h=1, 2, \cdots)$ can contain more than a single element $\neq 0$.

2. Let now $|w|>|t|>1$. We consider the function

$$G(w, t) = \Delta(w)g'(w)^{1/2}g'(t)^{1/2}/\{\Delta(t)(g(t) - g(w))\}. \tag{5}$$

It is regular for t fixed in $|w|>|t|$ and has a simple zero at infinity. Therefore it admits of a representation of the form

$$G(w, t) = \phi_1(t)w^{-1} + \phi_2(t)w^{-2} + \phi_3(t)w^{-3} + \cdots. \tag{6}$$

We shall show that

$$(2\pi i)^{-1}\lambda_k \int_{|t|=r} t^k G(w, t)dt = Q_k(w), \quad |w| > r > 1. \tag{7}$$

Indeed the left-hand member is, on account of (4),

$$(2\pi i)^{-1}\Delta(w)g'(w)^{1/2} \int_{|t|=r} \frac{p_k[g(t)]}{g(t) - g(w)} g'(t)dt - (2\pi i)^{-1} \int_{|t|=r} Q_k(t)G(w, t)dt.$$

On writing $\tau = g(t)$ we obtain for the first term

$$(2\pi i)^{-1}\Delta(w)g'(w)^{1/2} \int_{C_r} \frac{p_k(\tau)}{\tau - z} d\tau = 0$$

since $z=g(w)$ is outside C_r. The integral of the second term, being taken over a large circle $|t|=R$, tends to 0 as $R\rightarrow\infty$. Thus we get the residue $Q_k(w)$.

An alternative form of this result is

$$(2\pi i)^{-1}\lambda_k \int_{|t|=r} t^k\phi_h(t)dt = \lambda_{kh} \quad (k = 0, 1, 2, \cdots; h = 1, 2, 3, \cdots). \tag{8}$$

3. There is no difficulty in obtaining explicit representations for the functions $\phi_h(t)$. From (6) we have

$$\begin{aligned} \phi_1(t) &= -\Delta(\infty)g^{-1/2}g'(t)^{1/2}/\Delta(t), \\ \phi_2(t) &= -\Delta(\infty)g^{-1/2}(g'(t)^{1/2}/\Delta(t))(g(t)/g + \text{const.}). \end{aligned} \tag{9}$$

A direct expansion shows that $\phi_h(t)$ is of the form $g'(t)^{1/2}/\Delta(t)$ multiplied by a polynomial in $g(t)$ of the exact degree $(h-1)$.

In virtue of (8) and of the remark above concerning the vanishing of the λ_{kh} in a fixed column, we see at once that the Laurent series expansion of $\phi_h(t)$ cannot involve more than one negative power of t, that is, $\phi_h(t)$ must be of the form

$$bt^{-\beta} + b_0 + b_1 t + \cdots + b_{h-1}t^{h-1}, \quad b_{h-1} \neq 0, \quad \beta > 0.$$

As a consequence of this $\phi_2(t)/\phi_1(t)$, hence also $g(t)$, must be rational. This function cannot have other poles than 0 and ∞; otherwise $\phi_h(t)$ would have a further pole provided h is sufficiently large. Thus we find

$$g(t) = gt + g_0 + g_1 t^{-1} + \cdots + g_\sigma t^{-\sigma}. \tag{10}$$

4. On denoting the exact orders of $\phi_1(t)$ and of $g(t)$ at $t=0$ by ρ and σ respectively we first assume $\sigma = 0$, that is,

$$g(t) = gt + g_0, \quad \phi_1(t) = -\Delta(\infty)/\Delta(t) = bt^{-\rho} + b_0.$$

This yields types I and II given in the Introduction.

Next assume $\sigma > 0$. We now distinguish two principal cases.

(a) $Q_0(w)$ is not identically zero, that is, there is at least one coefficient $\lambda_{0h} \neq 0$. We know by (8) that $\phi_h(t)$ has a simple pole at $t=0$. Then, by (9),

$$\rho + (h-1)\sigma = 1.$$

Consequently we have to consider the following possibilities:

$$h = 1, \quad \rho = 1;$$
$$h = 2, \quad \sigma = 1, \quad \rho = 0.$$

Under the first hypothesis we have on account of (4) for $k=0$,

$$g'(t)^{1/2}/\Delta(t) = b_0 + b_1 t^{-1}, \quad g'(t)^{1/2}\Delta(t) = c_0 + c_1 t^{-1}, \quad b_0, b_1, c_0, c_1 \neq 0,$$

whence

$$g'(t) = b_0 c_0 + (b_0 c_1 + b_1 c_0)t^{-1} + b_1 c_1 t^{-2},$$

so that $b_0 c_1 + b_1 c_0 = 0$ and

$$g(t) = b_0 c_0 t + g_0 - b_1 c_1 t^{-1} = g_0 + b_0 c_0 \left(t + \frac{b_1^2}{b_0^2} t^{-1}\right),$$

while

$$\Delta(t) = \left[b_0 c_0 \left(1 - \frac{b_1^2}{b_0^2} t^{-2}\right)\right]^{1/2} / (b_0 + b_1 t^{-1}) = \left(\frac{c_0}{b_0}\right)^{1/2} \left(\frac{b_0 - b_1 t^{-1}}{b_0 + b_1 t^{-1}}\right)^{1/2}.$$

This is our type V.

The second hypothesis gives at once

$$g(t) = gt + g_0 + g_1 t^{-1} \quad (g, g_1 \neq 0); \quad g'(t)^{1/2}/\Delta(t) = \text{const.},$$

which is type IV.

(b) $Q_0(w)$ is identically zero, that is, $\Delta(t)g'(t)^{1/2} = \text{const}$. Then

$$\phi_1(t) = \text{const. } g'(t), \quad \rho = \sigma + 1,$$

and from (4) for $k=1$ we find that $\lambda_{1\sigma} \neq 0$. Consequently $\phi_\sigma(t)$ has a pole at $t=0$ of the exact order 2. Now the exact order of the pole $t=0$ of $\phi_h(t)$ is

$$\rho + (h-1)\sigma = \sigma + 1 + (h-1)\sigma = h\sigma + 1.$$

For $h=\sigma$ we have $\sigma^2+1=2$, $\sigma=1$, which corresponds to type III. Our proof is now complete.

KÖNIGSBERG, PR.

CERTAIN CONTACT PROPERTIES OF LINEAR SYSTEMS OF HYPERSURFACES*

BY

B. C. WONG

Introduction. The equation

$$\xi_0 f^{(0)} + \xi_1 f^{(1)} + \cdots + \xi_\rho f^{(\rho)} = 0, \tag{1}$$

where the ξ's are homogeneous parameters and the f's are homogeneous functions of degree n in the variables $x_0, x_1, \cdots, x_r$, represents an ∞^ρ-system, $|V|$, of hypersurfaces of order n in an r-space, S_r. Interpreting the ξ's as the homogeneous coordinates of a point of a ρ-space, Σ_ρ, we have a one-to-one correspondence between the points of Σ_ρ and the hypersurfaces of $|V|$.

In this paper we propose to deal with certain contact properties of the system $|V|$ and to investigate some of the properties and relations to one another of the corresponding loci in Σ_ρ. Our results will be generalizations of certain results which W. L. Edge† has, in connection with his study of octadic surfaces, described for the case $r=3$, $\rho=2$, $n=2$. We should mention that T. R. Hollcroft has derived‡ the properties of the curve in the plane Σ_2 of the parameters which corresponds to the Jacobian curve of a net of hypersurfaces in S_r, and in another paper§ those of the surface in the 3-dimensional space Σ_3 of the parameters which corresponds to the Jacobian surface of a web of surfaces in S_3. In neither of these papers, however, has the author touched upon the results which we are going to derive.

In the following we shall, first, describe the general case. This is done in §1. Then we shall consider the case $r=\rho=2$ in §2 and the case $n=2$, $r=\rho$ general in §3. Finally, in §4, we shall conclude the paper with a description of the results for $n=2$, $r=\rho=3$.

1. **The general case.** Let n, r, ρ be general. Since for a hypersurface to acquire a conical point or hypernode is equivalent to one condition, there are $\infty^{\rho-1}$ hypersurfaces of the system $|V|$ each possessing a hypernode. The

* Presented to the Society, June 20, 1934, under a slightly different title; received by the editors May 2, 1934.

† *Octadic surfaces and plane quartic curves*, Proceedings of the London Mathematical Society, (2), vol. 34 (1932), pp. 492–525.

‡ *Nets of manifolds in i dimensions*, Annali di Matematica, (4), vol. 5 (1927–1928), pp. 261–267.

§ *The general web of algebraic surfaces of order n and the involution defined by it*, these Transactions, vol. 35 (1933), pp. 855–868.

locus of these hypernodes is a $(\rho-1)$-dimensional manifold, the Jacobian manifold, $J^m_{\rho-1}$, of $|V|$. Differentiating equation (1) partially with respect to x_i, we have, writing f_i for $\partial f/\partial x_i$,

$$(2) \qquad \xi_0 f_i^{(0)} + \xi_1 f_i^{(1)} + \cdots + \xi_\rho f_i^{(\rho)} = 0 \qquad [i = 0, 1, \cdots, r].$$

The result of eliminating the ξ's from these equations is

$$(3) \qquad \left\| \begin{matrix} f_0^{(0)} & f_1^{(1)} & \cdots & f_r^{(r)} \\ f_0^{(1)} & f_1^{(1)} & \cdots & f_r^{(1)} \\ \cdot & \cdot & \cdot & \cdot \\ f_0^{(\rho)} & f_1^{(\rho)} & \cdots & f_r^{(\rho)} \end{matrix} \right\| = 0,$$

which are the equations of $J^m_{\rho-1}$. To determine m, notice that the matrix in (3) contains $\rho+1$ rows and $r+1$ columns. Then the system of equations given by (3) is a restricted system of equations and its order is, the elements of the matrix being all of order $n-1$,*

$$m = \binom{r+1}{r-\rho+1}(n-1)^{r-\rho+1}.$$

Now if we eliminate the x's from (2), we obtain a result in the form

$$\Delta(\xi_0, \xi_1, \cdots, \xi_\rho) = 0$$

which is of degree†

$$\mu = (r+1)(n-1)^r$$

and therefore represents a hypersurface $\Delta^\mu_{\rho-1}$ of order μ in the ρ-space Σ_ρ. The points of this hypersurface give the hypersurfaces of $|V|$ that have each a hypernode and may be said to correspond one-to-one to the points of $J^m_{\rho-1}$.

Let us now consider a variety V_k^N of k dimensions in S_r and let it be the complete intersection of $r-k$ hypersurfaces of orders $n_1, n_2, \cdots, n_{r-k}$ given by the equations

$$(4) \qquad F^{(1)} = 0, F^{(2)} = 0, \cdots, F^{(r-k)} = 0$$

respectively. Then $N = n_1 n_2 \cdots n_{r-k}$. Contact being one condition, there are $\infty^{\rho-1}$ hypersurfaces of $|V|$ tangent to V_k^N and the locus of the points of con-

* Salmon, *Modern Higher Algebra*, 4th edition, Lesson 19.

† The method of deriving this result is analogous to or an extension of the one described by Salmon in his *Analytic Geometry of Three Dimensions*, 5th edition, vol. 2, Art. 576.

tact is V_k^N itself if $k<\rho$ and a $V_{\rho-1}^{N'}$ if $k\geqq\rho$. The equations of $V_{\rho-1}^{N'}$ are the equations (4) and

$$M \equiv \left\| \begin{array}{cccc} f_0^{(0)} & f_1^{(0)} & \cdots & f_r^{(0)} \\ f_0^{(1)} & f_1^{(1)} & \cdots & f_r^{(1)} \\ \cdot & \cdot & \cdot & \cdot \\ f_0^{(\rho)} & f_1^{(\rho)} & \cdots & f_r^{(\rho)} \\ F_0^{(1)} & F_1^{(1)} & \cdots & F_r^{(1)} \\ \cdot & \cdot & \cdot & \cdot \\ F_0^{(r-k)} & F_1^{(r-k)} & \cdots & F_r^{(r-k)} \end{array} \right\| = 0,$$

that is, $V_{\rho-1}^{N'}$ is the intersection of V_k^N and the Jacobian manifold of the hypersurfaces given by (4) and any $\rho+1$ independent hypersurfaces of $|V|$. In the above we have written F_i for $\partial F/\partial x_i$.

Now to determine the order N' of $V_{\rho-1}^{N'}$, it is necessary first to determine the order of the restricted system of equations $M=0$. We see that M contains $r-k+\rho+1$ rows and $r+1$ columns and that the elements f_i are all of order $n-1$ and $F_i^{(j)}$ are of order n_j-1. Therefore, the order of M is*

$$\sum_{i=0}^{k-\rho+1} C_{k-\rho+1-i}H_i,$$

where

$$C_0 = H_0 = 1,$$

$$C_{k-\rho+1-i} = \binom{r+1}{k-\rho+1-i}(n-1)^{k-\rho+1-i},$$

$$H_i = \sum_q \sum (n_1-n)^{q_1}(n_2-n)^{q_2}\cdots(n_i-n)^{q_i}$$

where

$$q = q_1 + q_2 + \cdots + q_i = i.$$

Therefore, the order of $V_{\rho-1}^{N'}$ is N times the order of M, that is,

$$N' = n_1 n_2 \cdots n_{r-k} \sum_{i=0}^{k-\rho+1} C_{k-\rho+1-i}H_i.$$

Correspondingly, we find that there is a hypersurface $\Phi_{\rho-1}^{\nu}$ of order ν in Σ_ρ whose points yield the hypersurfaces of $|V|$ tangent to V_k^N and may be

* Salmon, *Modern Higher Algebra*, 4th edition, Lesson 19.

said to correspond to the points of $V_{\rho-1}^{N'}$. The condition of contact between a hypersurface of $|V|$ and V_k^N is the vanishing of the tact-invariant which involves the coefficients in (1) in the degree ν. We have, then,*

$$\nu = n_1 n_2 \cdots n_{r-k} \sum_{i=0}^{k} \binom{r+1}{k-i} H_i. \tag{5}$$

For $k<\rho$, the hypersurface $\Phi_{\rho-1}^{\nu}$ is the locus of $\infty^k(\rho-k-1)$-spaces and, for $k \geqq \rho$, it is a point locus. As our knowledge of hypersurfaces is still very scanty, we find it difficult to describe the $\Phi_{\rho-1}^{\nu}$'s for all values of $k<r$ and their relations to one another. We shall, however, content ourselves with a few special cases.

2. **The case $r=\rho=2$.** Now we have a net $|c|$ of curves of order n in an x-plane to which correspond the points of a ξ-plane. The Jacobian curve J^m of $|c|$ is of order $m=3(n-1)$ and the corresponding curve Δ^μ in the ξ-plane is of order $\mu=3(n-1)^2$. The characteristics of this curve Δ^μ are known.†

Now consider a curve C^{n_1} of order n_1 in the x-plane. Here $k=1$. From result (5) we have $\nu=n_1(n_1+2n-3)$ for the order of the curve Γ^ν in the ξ-plane whose points give the curves of the net tangent to C^{n_1}. If C^{n_1} has d_1 nodes and k_1 cusps, the quantity $2d_1+3k_1$ is to be deducted from the above result. Since the points of Γ^ν and the points of C^{n_1} are in one-to-one correspondence, the genus of Γ^ν is

$$\begin{aligned} p &= \tfrac{1}{2}(\nu-1)(\nu-2) - \delta - \kappa \\ &= \tfrac{1}{2}(n_1-1)(n_1-2) - d_1 - k_1, \end{aligned}$$

where δ is the number of nodes on Γ^ν and is equal to the number of the curves of $|c|$ doubly tangent to C^{n_1}, and κ is the number of cusps on Γ^ν and is equal to the number of the curves of $|c|$ having each a 3-point contact with C^{n_1}. Hence, the characteristics of Γ^ν are

$$\begin{aligned} \nu &= n_1(n_1+2n-3) - 2d_1 - 3k_1, \ddagger \\ \delta &= \tfrac{1}{2}\nu(\nu-9) - \tfrac{1}{2}n_1(n_1-9) + 3n_1(n-1) + d_1, \\ \kappa &= 3n_1(n+n_1-3) - 6d_1 - 8k_1, \\ \nu' &= n_1 n, \end{aligned}$$

the last being the class of the curve.

---

* Salmon, *Analytic Geometry of Three Dimensions*, above.

† See Hollcroft, *Nets of manifolds in i dimensions*, above.

‡ This formula is given in Salmon, *Higher Plane Curves*, 3d edition, Art. 97.

Since the given curve C^{n_1} meets J^m in $n_1m=3n_1(n-1)$ points, the curve Γ^ν touches the corresponding curve $3n_1(n-1)$ times and meets it again in $\mu\nu-6n_1(n-1)=3n_1(n-1)(n_1n+2n^2-5n-n_1+1)$ points. This is the number of the curves of $|c|$ which possess each a node and are tangent to C^{n_1}. Γ^ν touching Δ^μ $3n_1(n-1)$ times and possessing δ nodes and κ cusps must satisfy $3n_1(n-1)+\delta+2\kappa$ conditions and the degree of its freedom is $\frac{1}{2}\nu(\nu+3)-3n_1(n-1)-\delta-2\kappa$ which is equal to $\frac{1}{2}n_1(n_1+3)-d_1-2k_1$, the degree of freedom of C^{n_1} in the x-plane. Hence, we have a one-to-one correspondence between the curves of the type Γ^ν in the ξ-plane and all the curves of order n_1 for all values of n_1 in the x-plane.

For $n=r=\rho=2$, we have a net of conics in a plane, the x-plane. Corresponding to the conics of this net are the points of another plane, the ξ-plane. In the x-plane there is the Jacobian curve, J^3, the locus of the vertices of the degenerate conics of the net, and in the ξ-plane there is a cubic curve, Δ^3, whose points correspond to the degenerate conics of the net or may be said to correspond to the points of the Jacobian J^3. The conics of the net which are tangent to a given line l of the x-plane are given by the points of a conic λ^2 in the ξ-plane. This conic λ^2 is triply tangent to Δ^3, the points of contact being the images of the points on J^3 in which l intersects it. Thus, we have a one-to-one correspondence between the lines of the x-plane and the conics triply tangent to Δ^3 in the ξ-plane. Now through a given point A of the x-plane pass the conics of a pencil belonging to the net and in the ξ-plane we have a line α. Then, a one-to-one correspondence exists between the points of the x-plane and the lines of the ξ-plane. If A is on l, then α is tangent to λ^2 at the point which corresponds to the conic of the net tangent to l at A. We may look for loci of points in the ξ-plane which yield conics tangent to given curves in the x-plane and describe their properties and their relations to one another and to the curve Δ^3. But we shall not unnecessarily lengthen our work by a discussion of these details here as they can be easily obtained from the results above.

3. **The case $n=2$, $r=\rho$ general.** In this case we have an ∞^r-system of $|Q|$ of quadric hypersurfaces in an r-space S_r and correspondingly the points of another r-space Σ_r. The Jacobian of $|Q|$ is a hypersurface J_{r-1}^{r+1} of order $r+1$ in S_r and the locus of points in Σ_r corresponding to J_{r-1}^{r+1} is a hypersurface Δ_{r-1}^{r+1} also of order $r+1$.

Let now an $(r-t)$-space, S_{r-t}, be given in S_r and let its equation be

$$u_0^{(j)}x_0+u_1^{(j)}x_1+\cdots+u_r^{(j)}x_r=0 \qquad [j=1,2,\cdots,t].$$

The condition that a quadric hypersurface of $|Q|$ be tangent to S_{r-t} is

$$\begin{vmatrix} A_{00} & A_{01} & \cdots & A_{0r} & u_0^{(1)} & \cdots & u_0^{(t)} \\ A_{10} & A_{11} & \cdots & A_{1r} & u_1^{(1)} & \cdots & u_1^{(t)} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ A_{r0} & A_{r1} & \cdots & A_{rr} & u_r^{(1)} & \cdots & u_r^{(t)} \\ u_0^{(1)} & u_1^{(1)} & \cdots & u_r^{(1)} & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ u_0^{(t)} & u_1^{(t)} & \cdots & u_r^{(t)} & 0 & \cdots & 0 \end{vmatrix} = 0,$$

where $A_{ij}=a_{ij}^{(0)}\xi_0+a_{ij}^{(1)}\xi_1+\cdots+a_{ij}^{(r)}\xi_r$, $a_{ij}^{(h)}$ being the coefficients in the equation $f^{(h)}=0$ of the hth quadric hypersurface determining the system $|Q|$. The equation above is that of a hypersurface Φ_{r-1}^{r-t+1} of order $r-t+1$ in Σ_r whose points correspond to the quadric hypersurfaces of $|Q|$ tangent to S_{r-t}. We may say that there is a one-to-one correspondence between the points of S_{r-t} and the points of Φ_{r-1}^{r-t+1}.

Suppose $t=1$. Then, to all the ∞^{r+1} hypersurfaces of S_r correspond ∞^{r+1} hypersurfaces of the type Φ_{r-1}^{r} in Σ_r. Each such Φ_{r-1}^{r} is tangent to Δ_{r-1}^{r+1} along an $(r-2)$-dimensional variety, $\Theta_{r-2}^{r(r+1)/2}$, of order $r(r+1)/2$, and the points of this contact variety correspond to the points of the section of J_{r-1}^{r+1} by the corresponding hyperplane S_{r-1}.

Now suppose $t=2$ and let S_{r-2} be in a fixed S_{r-1}. The locus of points in Σ_r corresponding to those quadric hypersurfaces of $|Q|$ tangent to S_{r-2} is a hypersurface Φ_{r-1}^{r-1} of order $r-1$. This Φ_{r-1}^{r-1} is tangent to Δ_{r-1}^{r} corresponding to the fixed S_{r-1} along an $(r-2)$-dimensional variety $\Theta_{r-2}^{r(r-1)/2}$ and tangent to Δ_{r-1}^{r+1} along an $(r-3)$-dimensional variety $\Theta_{r-3}^{r(r+1)(r-1)/6}$. The points of $\Theta_{r-2}^{r(r-1)/2}$ correspond to the quadric hypersurfaces of $|Q|$ tangent to S_{r-1} at the points of S_{r-2}. The points of S_{r-2} considered as points of S_{r-1} have for corresponding points the points of $\Theta_{r-2}^{r(r-1)/2}$ lying on Φ_{r-1}^{r}. The points of S_{r-2} considered as points of contact between S_{r-2} and those quadric hypersurfaces of $|Q|$ which are tangent to it without being tangent to S_{r-1} have for corresponding points the points of Φ_{r-1}^{r-1}. Now S_{r-2} meets J_{r-1}^{r+1} in a V_{r-3}^{r+1} to which correspond the points of $\Theta_{r-3}^{r(r+1)(r-1)/6}$.

Let us now consider a series of sub-spaces $S_{r-1}, S_{r-2}, \cdots, S_{r-t}, \cdots, S_2, S_1, S_0$, each being contained in the preceding one. Then we have a series of hypersurfaces in Σ_r,

$$\Phi_{r-1}^{r},\ \Phi_{r-1}^{r-1},\ \cdots,\ \Phi_{r-1}^{r-t+1},\ \cdots,\ \Phi_{r-1}^{3},\ \Phi_{r-1}^{2},\ \Phi_{r-1}^{1}$$

[the last one being a hyperplane] corresponding to the given sub-spaces respectively. We see that

$$\begin{array}{llllllll}
\Phi_{r-1}^{r} & \text{is tangent to} & & & \Phi_{r-1}^{r+1} & \text{along a} & & \Theta_{r-2}^{r(r+1)/2}, \\
\Phi_{r-1}^{r-1} & `` & `` & `` & \Phi_{r-1}^{r} & `` & `` & \Theta_{r-2}^{r(r-1)/2}, \\
\cdot\;\cdot\;\cdot & & & & \cdot\;\cdot\;\cdot & & & \cdot\;\cdot\;\cdot \\
\Phi_{r-1}^{r-t+1} & `` & `` & `` & \Phi_{r-1}^{r-t+2} & `` & `` & \Theta_{r-2}^{(r-t+2)(r-t+1)/2}, \\
\cdot\;\cdot\;\cdot & & & & \cdot\;\cdot\;\cdot & & & \cdot\;\cdot\;\cdot \\
\Phi_{r-1}^{2} & `` & `` & `` & \Phi_{r-1}^{3} & `` & `` & \Theta_{r-2}^{3}, \\
\Phi_{r-1}^{1} & `` & `` & `` & \Phi_{r-1}^{2} & `` & `` & \Theta_{r-2}^{1}.
\end{array}$$

A general plane in Σ_r meets Δ_{r-1}^{r+1} and all the Φ's in a series of curves, an $(r+1)$-ic, an r-ic, $\cdots$, a cubic, a conic, and a line. The $(r+1)$-ic is touched by the r-ic at $r(r+1)/2$ points, the r-ic by the $(r-1)$-ic at $r(r-1)/2$ points, $\cdots$, the cubic by the conic at 3 points, and, finally, the conic is touched by the line. All these different groups of points of contact are the intersections of the plane and the various Θ_{r-2}'s.

Now

$$\begin{array}{llllllll}
\Theta_{r-2}^{r(r-1)/2} & \text{is tangent to} & & & \Theta_{r-2}^{r(r+1)/2} & \text{along a} & & \Theta_{r-3}^{r(r+1)(r-1)/6}, \\
\Theta_{r-2}^{(r-1)(r-2)/2} & `` & `` & `` & \Theta_{r-2}^{r(r-1)/2} & `` & `` & \Theta_{r-3}^{r(r-1)(r-2)/6}, \\
\cdot\;\cdot\;\cdot & & & & \cdot\;\cdot\;\cdot & & & \cdot\;\cdot\;\cdot \\
\Theta_{r-2}^{3} & `` & `` & `` & \Theta_{r-2}^{6} & `` & `` & \Theta_{r-3}^{4}, \\
\Theta_{r-2}^{1} & `` & `` & `` & \Theta_{r-2}^{3} & `` & `` & \Theta_{r-3}^{1}.
\end{array}$$

A general 3-space in Σ_r meets the whole configuration consisting of Δ_{r-1}^{r+1} and the Φ's in a configuration consisting of a series of surfaces F^{r+1}, F^r, $\cdots$, F^3, F^2, and a plane. The plane is tangent to F^2 which is a quadric cone along a line; F^2 is tangent to F^3 which is a four-nodal cubic surface along a twisted cubic curve C^3; F^3 is tangent to F^4, a quartic symmetroid known to have ten conical points, along a sextic curve C^6 of genus 3; $\cdots$; F^r is tangent to F^{r+1} along a curve $C^{r(r+1)/2}$. These curves of contact are the intersections of the Θ_{r-2}'s by the given 3-space. Now the line of contact between the plane and F^2 is tangent to C^3 at one point, C^3 is tangent to C^6 at 4 points, C^6 is tangent to C^{10} at 10 points, $\cdots$, $C^{r(r-1)/2}$ is tangent to $C^{r(r+1)/2}$ at $r(r+1)(r-1)/6$ points. These points of contact are the points where the given 3-space meets the various Θ_{r-3}'s.

We may continue this process until it terminates. But instead we shall exhibit the series of Φ's and the different series of Θ's in the following table, where for the moment (to simplify typography) we have set

$$(m, n) = \binom{m}{n}.$$

Sub-spaces in S_r	Corre-sponding Φ's in Σ_r	Varieties of contact				
J_{r-1}^{r+1}	Δ_{r-1}^{r+1}					
		$\Theta_{r-2}^{(r+1,2)}$				
S_{r-1}	Φ_{r-1}^{r}		$\Theta_{r-3}^{(r+1,3)}$			
		$\Theta_{r-2}^{(r,2)}$		$\Theta_{r-4}^{(r+1,4)}$		
S_{r-2}	Φ_{r-1}^{r-1}		$\Theta_{r-3}^{(r,3)}$		$\Theta_{r-5}^{(r+1,5)}$	
		$\Theta_{r-2}^{(r-1,2)}$		$\Theta_{r-4}^{(r,4)}$		.
S_{-3}	Φ_{r-1}^{r-2}		$\Theta_{r-3}^{(r-1,3)}$		.	
		$\Theta_{r-2}^{(r-2,2)}$		.		.
S_{r-4}	Φ_{r-1}^{r-3}		.		.	
. . .	. . .	. . .	. . .	. . .	. . .	. . .
S_2	Φ_{r-1}^{3}		Θ_{r-3}^{4}		.	
		Θ_{r-2}^{3}		Θ_{r-4}^{1}		
S_1	Φ_{r-1}^{2}		Θ_{r-3}^{1}			
		Θ_{r-2}^{1}				
S_0	Φ_{r-1}^{1}					

In this table any two symbols in the same column (other than the first) represent two varieties which touch along a variety whose symbol stands in the two diagonals in which the two symbols of the touching varieties lie. For example, Φ_{r-1}^{r} and Φ_{r-1}^{r-3} touch along $\Theta_{r-4}^{r(r-1)(r-2)(r-3)/24}$ along which $\Theta_{r-2}^{r(r-1)/2}$ and $\Theta_{r-2}^{(r-1)(r-2)/2}$, and also $\Theta_{r-3}^{r(r-1)(r-2)/6}$ and $\Theta_{r-3}^{(r-1)(r-2)(r-3)/6}$ touch.

4. **The case $n=2$, $r=\rho=3$.** We have now a web of quadric surfaces in S_3 whose Jacobian is a quartic surface J_2^4. The locus of points in Σ_3 corresponding to J_2^4 is a quartic symmetroid Δ_2^4. A plane f in S_3 meets J_2^4 in a quartic curve c^4. The points of Σ_3 corresponding to those quadric surfaces of the web tangent to f or to the points of f as points of contact form a four-nodal cubic surface F^3 which is tangent to Δ_2^4 along a sextic curve Θ_1^6. Θ_1^6 passes through the four conical points of F^3 and its points correspond to the points of c^4. We have in Σ_3 ∞^3 four-nodal cubic surfaces each having its four nodes lying on Δ_2^4 and touching Δ_2^4 along a sextic curve of genus 3.

Now take a line l in f. The locus of points in Σ_3 corresponding to the quadric surfaces of the web tangent to l or to the points of l as points of contact is a quadric cone F^2. The vertex of F^2 corresponds to the quadric surface of the web passing through l and it lies on F^3 corresponding to f. The line l considered as a line of f has for image a cubic curve Θ_1^3 on F^3. F^2 and F^3 touch along Θ_1^3. Now Θ_1^3 and the curve Θ_1^6 corresponding to c^4 common to J_2^4 and f are tangent at 4 points at which F^2 and Δ_2^4 are also tangent. These four points correspond to the four points common to c^4 and l in f. F^2 intersects Δ_2^4 in an octavic curve with four actual double points and the points of this curve give the cones of the web tangent to l. The locus of the vertices of these cones is a curve of order 12 on J_2^4 and it also has four actual double points. We see that there are ∞^2 quadric cones in Σ_3 corresponding to the ∞^2 lines of f each of which has its vertex on F^3 and is tangent to F^3 along a twisted cubic curve passing through the four nodes of F^3. Each point of F^3 is the vertex of two such cones, for a quadric surface of the web containing a line of f contains another line of f.

If we take any curve, say a conic, c^2, in f, then the quadric surfaces of the web tangent to c^2 are given by the points of a ruled sextic surface F^6 in Σ_3. This surface F^6 is, in fact, a developable surface and it is tangent to F^3 along a sextic curve γ^6 whose points yield the quadric surfaces of the web tangent to f with the points of contact on c^2. Its edge of regression is also a sextic curve and the nodal curve is a quartic, the points on the former giving the quadric surfaces which osculate c^2 and the points on the latter giving those doubly tangent to c^2. If c^2 happens to be on a quadric surface Q of the web, F^6 is a tangent cone of F^3 having four nodal and six cuspidal elements and its vertex corresponds to Q. In either case, F^6 is tangent to Δ_2^4 at 8 points at which γ^6 is tangent to Θ_1^6, the curve of contact between F^3 and Δ_2^4. These 8 points correspond to the 8 points in which c^2 meets J_2^4. Now the two surfaces, F^6 and F^3, intersect in another sextic curve and the points of this curve yield the quadric surfaces of the web which are tangent both to f and to c^2.

Finally, let a point P be given in f. The quadric surfaces of the web passing through P form a net and they are given by the points of a plane π in Σ_3. If P is on a fixed line l of f, π is tangent to F^2 along a line which is tangent to Θ_1^3 at one point. π is tangent to F^3 at the same point. There are ∞^1 cones of the web passing through P and to these cones correspond the points of the quartic curve in which π meets Δ_2^4. The vertices of these cones lie on a sextic curve of J_2^4.

University of California,
Berkeley, Calif.

A BASIS FOR RESIDUAL POLYNOMIALS IN n VARIABLES*

BY

MARIE LITZINGER

I. Introduction

Kempner† has established the existence of a basis for residual polynomials in one variable with respect to a composite modulus. A residual polynomial modulo m is by definition a polynomial $f(x)$ with integral coefficients which is divisible by m for every integral value of x, and a residual congruence is written $f(x)\equiv 0 \pmod{m}$. By a basis for a given modulus is meant a finite set of residual polynomials $p_i(x)$ which fulfills two requirements: (i) every residual polynomial modulo m is expressible as a sum of products of $p_i(x)$ by polynomials in x with integral coefficients; (ii) no member of the set $p_i(x)$ can be written identically equal to a sum of products of the remaining members of the set by polynomials in x with integral coefficients.

For this work, the following notation is used. The symbol $\mu(d)$ denotes the least positive integer for which d divides $\mu!$. A special set of divisors of m is chosen: separate all divisors of m which exceed 1 into groups such that $\mu(d)$ has the same value for all the d's of a group but different values for the d's of different groups; select the largest d of each group and denote this set by $d_1, \cdots, d_s$. Finally, $\Pi(\mu)=x(x-1)\cdots(x-\mu+1)$; when x is replaced by x_j, the product will be designated by $\Pi_j(\mu)$; $\Pi(1)$ is interpreted as 1. Employing this notation, Dickson‡ gave a brief proof of the theorem due to Kempner§:

Every residual polynomial $f(x)$ modulo m is a sum of products of m and $(m/d_i)\Pi(\mu(d_i))$ for $i=1, \cdots, s$ by polynomials in x with integral coefficients.

In a later paper, Kempner¶ considered the problem for n variables. In attempting to apply Dickson's method to the proof of the existence of a basis for residual polynomials in more than one variable, I found that Kempner had omitted from the set $p_i(x_1, \cdots, x_n)$ certain residual polynomials in

* Presented to the Society, February 23, 1935; received by the editors July 8, 1934.

† These Transactions, vol. 22 (1921), pp. 240–266.

‡ L. E. Dickson, *Introduction to the Theory of Numbers*, p. 26, Theorem 28.

§ These Transactions, vol. 22 (1921), p. 263.

¶ These Transactions, vol. 27 (1925), pp. 287–298.

several variables. This was brought to my attention by an example in two variables modulo 12. For this modulus, the $d_1, \cdots, d_s$ are $d_1=12$, $d_2=6$, $d_3=2$; the corresponding μ's are $\mu_1=4$, $\mu_2=3$, $\mu_3=2$. Write $q_i=m/d_i$. The part of the basis containing one variable is composed of

$$(1) \qquad 12, \quad q_i\Pi_1(\mu_i), \quad q_i\Pi_2(\mu_i) \qquad (i = 1, 2, 3).$$

Kempner would include in the basis $p_i(x_1, x_2)$ modulo 12 only the seven terms (1). However, the residual polynomial,

$$P = (m/(d_3 \cdot d_3))\Pi_1(\mu_3)\Pi_2(\mu_3) = 3x_1(x_1 - 1)x_2(x_2 - 1),$$

must be added since, as is shown below, it is impossible to write the identity

$$(2) \qquad P = 12 \cdot c + \sum_{i=1}^{3} q_i\Pi_1(\mu_i)f_i + \sum_{i=1}^{3} q_i\Pi_2(\mu_i)g_i,$$

where c, f_i, g_i are polynomials in x_1, x_2 with integral coefficients. By use of $(x_1, x_2)=(0, 0)$, $(2, 0)$, $(0, 2)$ we prove the constant terms of c, f_3, g_3 even. The pair $(x_1, x_2)=(2, 2)$ shows the right side of (2) divisible by 24 and the left side equal to 12.

II. Representation of residual polynomials

Dickson's method of establishing the existence of a basis for residual polynomials modulo m in one variable may be applied to the case of two variables and then by induction to n variables. Several preliminary steps are necessary. The first is the statement of two lemmas due to Dickson.

Lemma* 1. *If d is any divisor of m, $\mu(d)$ is the minimum degree of a residual polynomial $f(x)$ modulo m whose leading coefficient is m/d.*

Lemma† 2. *Any residual polynomial $f(x)$ modulo m is term by term congruent modulo m to the product of an integer prime to m by a residual polynomial whose leading coefficient is a divisor of m.*

The next step is to obtain a lemma similar to Lemma 2.

Lemma 3. *Any residual polynomial $f(x_1, \cdots, x_n)$ modulo m, written as a function of x_1 with coefficients containing $x_2, \cdots, x_n$, is term by term congruent modulo m to the product of an integer prime to m by a residual polynomial in which the greatest common divisor of the coefficients of the highest power of x_1 is a divisor of m.*

Let

$$f(x_1, \cdots, x_n) = cG(x_2, \cdots, x_n)x_1^r + \cdots,$$

* L. E. Dickson, *Introduction to the Theory of Numbers*, p. 25, V.

† Ibid., p. 25, VI.

where the coefficients of G have the greatest common divisor 1, and let g be the greatest common divisor of $c=gC$ and $m=gM$. Since C is prime to M, $CL\equiv 1 \pmod{M}$ has a unique solution L. Then every integer satisfying $Cz\equiv 1 \pmod{M}$ is of the form $z=L+My$, and y can be chosen so that z is prime to m. Consequently $zZ\equiv 1 \pmod{m}$ has a solution Z and $cz=gCz\equiv g \pmod{m}$, also

$$zf \equiv gG(x_2, \cdots, x_n)x_1^r + \cdots \pmod{m},$$

$$f \equiv Z[gG(x_2, \cdots, x_n)x_1^r + \cdots] \pmod{m}.$$

Finally, two properties of μ and divisors of m must be derived.

LEMMA 4. *If d and d' are divisors of m such that d' divides d, then $\mu(d')\leqq\mu(d)$.*

Write $d=p_1^{a_1}\cdot p_2^{a_2}\cdots p_k^{a_k}$ where $p_1, p_2, \cdots, p_k$ are distinct primes. Then $\mu(d)$ is the largest (or one of the largest in case several are equal)* of the numbers $\mu(p_1^{a_1}), \mu(p_2^{a_2}), \cdots, \mu(p_k^{a_k})$. Since d' divides d, $d'=p_1^{b_1}\cdot p_2^{b_2}\cdots p_k^{b_k}$ where $0\leqq b_i\leqq a_i$ for $i=1, 2, \cdots, k$. So $\mu(p_i^{b_i})\leqq\mu(p_i^{a_i})$ for $i=1, 2, \cdots, k$, and $\mu(d')$, the largest of the $\mu(p_1^{b_1}), \mu(p_2^{b_2}), \cdots, \mu(p_k^{b_k})$, is less than or equal to $\mu(d)$.

LEMMA 5. *If d_i is one of the set $d_1, \cdots, d_s$ for m, then d_i is divisible by every divisor of m which divides $\mu(d_i)!$.*

The assumption that a divisor d of m divides $\mu(d_i)!$ and does not divide d_i leads to a contradiction as follows. Denote by D the greatest common divisor of d_i and d so that $d_i=DD_i$ and $d=DD'$. Since $\mu(d_i)!$ is divisible by both d_i and d and since D_i and D' are relatively prime, $\mu(d_i)!$ is divisible by $N=DD_iD'$ and N divides m. As N divides $\mu(d_i)!$, $\mu(N)\leqq\mu(d_i)$; as N is divisible by d_i, by Lemma 4, $\mu(N)\geqq\mu(d_i)$; consequently $\mu(N)=\mu(d_i)$. But d_i is one of the set $d_1, \cdots, d_s$ and therefore is the maximum of all divisors d_j of m for which $\mu(d_j)=\mu(d_i)$. There is then a contradiction unless $D'=1$, therefore d divides d_i.

With the aid of these lemmas, it is possible to prove

THEOREM 1. *Every residual polynomial $f(x_1, x_2)$ modulo m is a sum of products of m and functions*

$$(m/(d_{i_1}\cdot d_{i_2}))\Pi_1(\mu(d_{i_1}))\Pi_2(\mu(d_{i_2})) \tag{3}$$

by polynomials in x_1, x_2 with integral coefficients, where d_{i_1}, d_{i_2} are divisors of m, at least one belongs to the set $d_1, \cdots, d_s$, and the product $d_{i_1}\cdot d_{i_2}$ divides m.

By Lemma 3,

---

* Kempner, these Transactions, vol. 22 (1921), p. 243.

$$f(x_1, x_2) = m\phi(x_1, x_2) + ZF(x_1, x_2),$$

where Z and the coefficients of ϕ are integers, Z is prime to m and F is a residual polynomial modulo m of the form

$$(m/d)G(x_2)x_1^r + \cdots, \tag{4}$$

d being a divisor of m. If $d=1$, the terms containing m as a factor may be combined with $m\phi$ and the remaining portion considered the new ZF. So let $d>1$.

Case 1. Let $r \geqq \mu(d)$. Employ the relation*

$$(m/d)\Pi_1(\mu(d)) = (qm/d_{i_1})\Pi_1(\mu(d_{i_1})), \tag{5}$$

where d_{i_1} is one of the set $d_1, \cdots, d_s$ and q is an integer. The product of (5) by $G(x_2)x_1^{r-\mu(d_{i_1})}$ gives a function whose term in x_1^r is identical with that of F. The difference is a residual polynomial of degree less than r in x_1.

Case 2. Let $r<\mu(d)$. Consider F which is of the form (4). For a chosen value x_2' of x_2, by Lemma 2, F as a residual polynomial in x_1 is term by term congruent modulo m to the product of an integer prime to m by a residual polynomial whose leading coefficient is a divisor of m, that is,

$$F(x_1, x_2') = (m/d)G(x_2')x_1^r + \cdots \equiv z((m/d')x_1^r + \cdots) \pmod{m},$$

where z is prime to m and d' divides m. Then $(m/d)G(x_2') = z\cdot m/d' + km$ where k is integral. As m/d' divides m, $(m/d)G(x_2')$ is divisible by m/d'. Now $(m/d')x_1^r + \cdots$ is a residual polynomial whose leading coefficient is a divisor of m, consequently, by Lemma 1, $r \geqq \mu(d')$. Let d_i represent the divisor of the set $d_1, \cdots, d_s$ to which corresponds the largest μ which does not exceed r. Then $\mu(d_i) \geqq \mu(d')$, therefore d' divides $\mu(d_i)!$. By Lemma 5, d' divides d_i. Consequently m/d_i divides m/d' and must then divide $(m/d)G(x_2')$.

There is an important consequence of the divisibility of $(m/d)G(x_2')$ by m/d_i. Note first that m/d_i does not divide m/d, for if d divides d_i, by Lemma 4, $\mu(d) \leqq \mu(d_i)$; but by the definition of d_i, $\mu(d_i) \leqq r$; the conclusion $\mu(d) \leqq r$ contradicts the hypothesis of this second case, namely $r<\mu(d)$. Since m/d_i does not divide m/d, denote their greatest common divisor by M. Then

$$m/d_i = Mg, \quad m/d = Mv, \quad v\cdot m/d_i = g\cdot m/d, \tag{6}$$

where $g>1$ and prime to v. From the divisibility of $(m/d)G(x_2')$ by m/d_i, it follows that the quotient of $(m/d)G(x_2')$ by m/d_i, which equals $(v/g)G(x_2')$, is integral. As g is prime to v, g divides $G(x_2')$.

Consider $G(x_2)$ for other values of x_2. Although the coefficient correspond-

* L. E. Dickson, *Introduction to the Theory of Numbers*, p. 27, equation (34).

ing to m/d' varies with the choice of x_2, d_i by definition is determined by r and m and is independent of the value of x_2. Consequently g and v, determined by m/d_i and m/d, do not vary with x_2. So for every choice of x_2, the m/d' determined by it is such that it divides $(m/d)G(x_2)$ and is divisible by m/d_i; therefore $(m/d)G(x_2)$ is divisible by m/d_i and $G(x_2)$ is divisible by g.

Since $G(x_2)$ is divisible by g for every value of x_2, $G(x_2)\equiv 0 \pmod{g}$. Therefore $G(x_2)$ is expressible* as a sum of products of g and $(g/d_{i_2})\Pi_2(\mu(d_{i_2}))$ by polynomials in x_2 with integral coefficients, where the d_{i_2} represent the set of divisors of g selected as the set $d_1, \cdots, d_s$ was chosen from all divisors of m. As g divides m, for each d_{i_2}, $\mu(d_{i_2})=\mu(d_h)$ where d_h is one of the set $d_1, \cdots, d_s$ for m and, by Lemma 5, d_{i_2} equals or divides d_h.

In the work which follows, write d_{i_1} for d_i to indicate its association with x_1. The term of F containing the highest power of x_1 may be expressed as follows:

$$(m/d)G(x_2)x_1^r = g(m/d)(1/g)G(x_2)x_1^r = v(m/d_{i_1})(t/m)G(x_2)x_1^r,$$

where t is defined by the equation $tg=m$, and g and v are defined by (6). Note that d_{i_1} divides t from the following considerations: $t/d_{i_1}=(m/g)g/(dv)$ $=m/(dv)$ which is integral since v divides m/d. As d_{i_1} divides t and each d_{i_2} is a factor of g, for every d_{i_2}, the product $d_{i_1}d_{i_2}$ divides m. From its definition, d_{i_1} is one of the set $d_1, \cdots, d_s$ for m. The product of $v(m/d_{i_1})\Pi_1(\mu(d_{i_1}))$ by $(t/m)G(x_2)x_1^{r-\mu(d_{i_1})}$ gives a function whose term in x_1^r is identical with that of F. The difference is a residual polynomial of degree less than r in x_1.

This process, continued for the resulting polynomials considered as functions of x_1 or x_2, lowers the degree in x_1 or x_2 at each step and leads to a difference zero. Finally $f(x_1, x_2)$ is expressed in the manner described in Theorem 1.

A similar theorem for n variables is readily proved by induction.

THEOREM 2. *Every residual polynomial $f(x_1, \cdots, x_n)$ modulo m is a sum of products of m and functions*

$$(m/(d_{i_1}\cdots d_{i_n}))\Pi_1(\mu(d_{i_1}))\cdots\Pi_n(\mu(d_{i_n})) \tag{7}$$

by polynomials in $x_1, \cdots, x_n$ with integral coefficients where the d_{i_j} are divisors of m, at least one of the $d_{i_1}, \cdots, d_{i_n}$ belongs to the set $d_1, \cdots, d_s$, and the product $d_{i_1}\cdots d_{i_n}$ divides m.

The theorem has been established for the case $n=2$. Assume that it holds for $n-1$ variables and show that it must then be true for n. By Lemma 3,

$$f(x_1, \cdots, x_n) = m\phi(x_1, \cdots, x_n) + ZF(x_1, \cdots, x_n),$$

where Z and the coefficients of ϕ are integers, Z is prime to m, and F is a resid-

* L. E. Dickson, *Introduction to the Theory of Numbers*, p. 26, Theorem 28.

ual polynomial modulo m of the form

$$(8)\qquad (m/d)G(x_2, \cdots, x_n)x_1^r + \cdots,$$

d being a divisor of m. If $d=1$, the terms containing m as a factor may be combined with $m\phi$ and the remaining portion considered the new ZF. So let $d>1$.

Case 1. Let $r \geqq \mu(d)$. Employ relation (5). The product of (5) by $G(x_2, \cdots, x_n)x_1^{r-\mu(d_i)}$ gives a function whose term in x_1^r is identical with that of F. The difference is a residual polynomial modulo m of degree less than r in x_1.

Case 2. Let $r<\mu(d)$. Consider F which is of the form (8). For a chosen set of values $x_2', \cdots, x_n'$, by Lemma 2, F as a residual polynomial in x_1 modulo m is term by term congruent modulo m to the product of an integer prime to m by a residual polynomial whose leading coefficient is a divisor of m, that is,

$$\begin{aligned} F(x_1, x_2', \cdots, x_n') &= (m/d)G(x_2', \cdots, x_n')x_1^r + \cdots \\ &\equiv z((m/d')x_1^r + \cdots) \qquad (\text{mod } m), \end{aligned}$$

where z is prime to m and d' divides m. For the chosen set $x_2', \cdots, x_n'$, $(m/d)G(x_2', \cdots, x_n') = z \cdot m/d' + km$ where k is integral. As m/d' divides m, m/d' divides $(m/d)G(x_2', \cdots, x_n')$. Now $(m/d')x_1^r + \cdots$ is a residual polynomial whose leading coefficient is a divisor of m, consequently, by Lemma 1, $r \geqq \mu(d')$.

Repeat the argument given in Theorem 1 for Case 2, defining d_i as the divisor of the set $d_1, \cdots, d_s$ for m to which corresponds the largest μ not exceeding r, and replacing the phrase "value of x_2'" by "set of values $x_2', \cdots, x_n'$," and $G(x_2')$ by $G(x_2', \cdots, x_n')$. Exactly as in the first two paragraphs of Case 2, Theorem 1, m/d_i divides m/d' and therefore divides $(m/d)G(x_2', \cdots, x_n')$; but m/d_i does not divide m/d. Let M denote the greatest common divisor of m/d_i and m/d, and obtain (6). Then the quotient of $(m/d)G(x_2', \cdots, x_n')$ by m/d_i, which equals $(v/g)G(x_2', \cdots, x_n')$, is integral. Since g is prime to v, g divides $G(x_2', \cdots, x_n')$.

Consider $G(x_2, \cdots, x_n)$ for other values of $x_2, \cdots, x_n$. Although the coefficient corresponding to m/d' varies with the choice of $x_2, \cdots, x_n$, d_i is determined by r and m and is independent of the values of $x_2, \cdots, x_n$. Consequently g and v, determined by m/d_i and m/d, do not vary with $x_2, \cdots, x_n$. So for every choice of $x_2, \cdots, x_n$, the m/d' determined by it is such that it divides $(m/d)G(x_2, \cdots, x_n)$ and is divisible by m/d_i; therefore $(m/d)G(x_2, \cdots, x_n)$ is divisible by m/d_i and $G(x_2, \cdots, x_n)$ is divisible by g.

Since $G(x_2, \cdots, x_n)$ is divisible by g for every set of values $x_2, \cdots, x_n$, $G(x_2, \cdots, x_n) \equiv 0$ (mod g). According to the hypothesis, G as a residual

polynomial in $n-1$ variables is expressible as a sum of products of g and

$$(g/(d_{i_2}\cdots d_{i_n}))\Pi_2(\mu(d_{i_2}))\cdots\Pi_n(\mu(d_{i_n}))$$

by polynomials in $x_2, \cdots, x_n$ with integral coefficients, where the d_{i_j} (for $j=2, \cdots, n$) represent divisors of g and $d_{i_2}\cdots d_{i_n}$ divides g. Since g divides m, for each d_{i_j}, $\mu(d_{i_j})=\mu(d_h)$ where d_h is one of the set $d_1, \cdots, d_s$ for m, and, by Lemma 5, d_{i_j} equals or divides d_h.

For the following work, write d_{i_1} in place of d_i to indicate its association with x_1. The term of F which contains the highest power of x_1 may be expressed as follows:

$$(m/d)G(x_2, \cdots, x_n)x_1^r = v(m/d_{i_1})(t/m)G(x_2, \cdots, x_n)x_1^r,$$

where t is defined by the equation $tg=m$, and g and v are defined by (6). As in the fifth paragraph of Case 2, Theorem 1, d_{i_1} divides t. Since each product $d_{i_2}\cdots d_{i_n}$ divides g, then $d_{i_1}d_{i_2}\cdots d_{i_n}$ divides m. The product of $v(m/d_{i_1})\cdot\Pi_1(\mu(d_{i_1}))$ by $(t/m)G(x_2, \cdots, x_n)x_1^{r-\mu(d_{i_1})}$ gives a function whose term in x_1^r is identical with that of F. The difference is a residual polynomial of degree less than r in x_1.

This process, continued for the resulting polynomials considered as functions of x_j for $j=1, 2, \cdots, n$, lowers the degree in x_j at each step and leads to a difference zero. Finally $f(x_1, \cdots, x_n)$ is expressed in the manner described in Theorem 2.

Theorems 1 and 2 contain one interesting difference from the theorem for one variable. Of the divisors of m appearing in a term (3) or (7), only one is necessarily chosen from $d_1, \cdots, d_s$.

III. Selection of a basis

It is now essential to establish a basis for residual polynomials in n variables modulo m. By Theorem 2, the set composed of m and all terms (7) fulfills the first requirement for a basis. It remains to select from m and (7) a reduced set $p_i(x_1, \cdots, x_n)$ such that no member of p_i can be written identically equal to a sum of products of the remaining p_i by polynomials in $x_1, \cdots, x_n$ with integral coefficients, and such that each of the terms of m and (7) not included among the p_i can be written identically equal to a sum of products of p_i by polynomials in $x_1, \cdots, x_n$ with integral coefficients. The terms p_i form a basis and will be called independent. All other terms m and (7) will be called dependent.

A term $k\cdot\Pi_1(\mu(d_{i_1}))\cdots\Pi_n(\mu(d_{i_n}))$ of (7) whose coefficient k is a multiple of that of another term (7) containing $\Pi_1(\mu(d_{i_1}))\cdots\Pi_n(\mu(d_{i_n}))$ is obviously dependent. Discard such terms and represent the remaining terms (7) by

$$(9) \qquad P(d_{i_1}, \cdots, d_{i_n}).$$

Denote by S the set composed of m and all terms (9). Throughout the discussion, an element of the set S will be termed simple or compound according as it contains one variable or more than one variable. The phrase "member related to (9)" will be used to designate any term of the set S, simple or compound, which contains not more than $\mu(d_{i_j})$ factors in x_j for each $j=1, \cdots, n$.

The following theorem establishes a basis.

THEOREM 3. *For the general modulus m, a basis for residual polynomials in n variables is composed of m, all simple terms and all compound terms* (9) *such that $\mu(d_{i_1}), \cdots, \mu(d_{i_n})$ are all multiples of the same prime factor of m.*

Example. For the modulus $3^3\cdot 5$, the set S in two variables contains terms which are not members of the basis. The $d_1, \cdots, d_s$ for this modulus are $d_1=3^3\cdot 5$, $d_2=3^2\cdot 5$, $d_3=3\cdot 5$, $d_4=3$; the corresponding μ's are $\mu_1=9$, $\mu_2=6$, $\mu_3=5$, $\mu_4=3$. The basis is composed of $3^3\cdot 5$, all simple terms, and the compound terms $P(d_4, d_4)$, $P(d_4, d_2)$, $P(d_2, d_4)$. The dependent compound terms of S are expressible in terms of the basis as follows:

$$P(d_4, d_3) = 2(x_2 - 3)(x_2 - 4)P(d_4, d_4) - 3x_1(x_1 - 1)(x_1 - 2)q_3\Pi_2(\mu_3),$$

$$P(d_3, d_4) = 2(x_1 - 3)(x_1 - 4)P(d_4, d_4) - 3x_2(x_2 - 1)(x_2 - 2)q_3\Pi_1(\mu_3).$$

The part of Theorem 3 concerned with simple terms is readily established. Kempner* proved that m and $(m/d_i)\Pi(\mu(d_i))$ for $i=1, \cdots, s$ form a basis for residual polynomials modulo m in one variable. It follows that m as well as each simple term of S is independent of all other members of the set S. For instance, to show the independence of a simple term in x_j, set the remaining $n-1$ variables equal to zero.

The proof of the portion of Theorem 3 which deals with compound terms will be divided into two parts. First it will be shown that each of the compound terms listed in the theorem is independent of all other members of the set S. Then it will be shown by means of an auxiliary theorem that these are the only independent compound terms.

It is not difficult to prove the independence of the compound terms described in Theorem 3. Suppose it were possible to write the identity

$$(10) \qquad P(d_{i_1}, \cdots, d_{i_n}) = m\cdot c + \sum_i P_i\cdot f_i,$$

where c and f_i are polynomials in $x_1, \cdots, x_n$ with integral coefficients and the P_i represent all members of the set S, simple and compound, except

$$(11) \qquad P(d_{i_1}, \cdots, d_{i_n}).$$

* These Transactions, vol. 22 (1921), pp. 263–264.

By hypothesis each μ on the left side of (10) is a multiple of a prime p which divides m. That each term on the right contains one more factor p than appears on the left for $x_1=\mu(d_{i_1}), \cdots, x_n=\mu(d_{i_n})$ may be shown as follows. Substitute successively for each x_j the values 0, $\mu(d_{i_k})$ for all $k=1, \cdots, n$ such that $\mu(d_{i_k})\leqq\mu(d_{i_j})$. Under this substitution, terms P_i not related to (11) disappear, and the constant term of c and each remaining f_i associated with a member of the basis containing only μ's which are multiples of p is proved congruent to zero modulo p. Related members not included among the latter present no difficulty since for the values listed above each will contain p to a power at least one greater than that exhibited in the modulus, p^s. For instance, consider the simple term $(m/d_i)\Pi((\mu(d_i))$ which lies between terms whose variable parts are $\Pi(rp)$ and $\Pi((r+1)p)$ and contain respectively rp and $(r+1)p$ factors. This implies $rp<\mu(d_i)<(r+1)p$. For $x=rp$, $(m/d_i)\cdot\Pi(\mu(d_i))$ is zero; for $x=\mu(d_i)$ it is divisible by exactly p^s. The sequence $\Pi(\mu(d_i))=x(x-1)\cdots(x-\mu(d_i)+1)$ contains at least one higher power of p for $x=(r+1)p$ than for $x=\mu(d_i)$ since one additional factor p is thus introduced at the beginning of the sequence when $(r+1)p$ is substituted for x, and none is lost at the end as the sequence contains more than rp factors. The same reasoning holds for members of the set S formed by compounding $(m/d_i)\Pi(\mu(d_i))$ with other terms; however it will be shown in the Auxiliary Theorem that such members are dependent. The independence of (11) follows immediately; substitute for each x_j the value $\mu(d_{i_j})$. The left side of (10) is divisible by exactly p^s, the right side by p^{s+1}.

It is possible to select from the set S certain dependent terms. If the coefficient of a compound member is the greatest common divisor of the coefficients of related terms, it is expressible as a linear homogeneous function of them with integral coefficients. Since the related terms by definition contain no more factors in any one variable than appear in the given term, the latter may be expressed by means of the related members in the manner described in Theorem 3.

That these are the only dependent terms follows from the

AUXILIARY THEOREM. *For a compound term* (11) *of the set S for n variables modulo m, if*

$$m/(d_{i_1}\cdots d_{i_n}) \tag{12}$$

is not the greatest common divisor of the coefficients of all related members of the set S, each $\mu(d_{i_j})$ for $j=1, \cdots, n$ is a multiple of the same prime factor of the modulus.

Adopt the following notation to designate terms related to (11): let d_{i_j}' represent any of the divisors of m such that $\mu(d_{i_j}')<\mu(d_{i_j})$ for $j=1, \cdots, n$.

Then the coefficient of any term related to (11) is of the form

$$m/(d_{k_1} \cdots d_{k_n}), \tag{13}$$

where at least one $d_{k_i} = d_{i_i}'$ if all of them are greater than 1; for one possible type of coefficients (13), each $d_{k_i} = d_{i_i}'$.

From the manner in which the set S is constructed, (12) divides the coefficients of all related terms. Since by hypothesis (12) is not the greatest common divisor of all coefficients (13), denote their greatest common divisor by the product of k by (12), where k is a divisor of m greater than 1 which may or may not be prime to (12). Then for $j=1, \cdots, n$, each m/d_{i_j}' contains a higher power of k than appears in m/d_{i_j}. Otherwise one of the combinations (13) would equal (12) and (12) would be the greatest common divisor of all coefficients of related terms. In other words, the coefficient of every simple term of the set S which exceeds m/d_{i_j} contains a higher power of k than does m/d_{i_j}. Since for each j, the largest $\Pi_j(\mu(d_{i_j}'))$ has a coefficient which divides all the other m/d_{i_j}', its coefficient contains the product of k by (12). So for each j, $\Pi_j(\mu(d_{i_j}))$ differs from the largest $\Pi_j(\mu(d_{i_j}'))$ in that it is divisible by a higher power of k than $\Pi_j(\mu(d_{i_j}'))$ for all values of x_j. Therefore if k is a prime p, for each j, $\mu(d_{i_j})$ is a multiple of p; if k is composite, for each j, $\mu(d_{i_j})$ is a multiple of the same prime factor of k.

There are two corollaries to Theorem 3.

COROLLARY 1. *For a modulus composed of the product of distinct primes, a basis is composed of m and all simple terms.*

Write the modulus m as $p_1 p_2 \cdots p_c$ where the p's are distinct primes arranged in descending order. The set $d_1, \cdots, d_s$ for m is composed of $p_1 p_2 p_3 \cdots p_c$, $p_2 p_3 \cdots p_c, \cdots, p_{c-1} p_c$, p_c; the corresponding μ's are p_1, $p_2, \cdots, p_{c-1}$, p_c. No product of two or more of the d's listed above will divide m since each contains p_c. Even when all possible divisors of m are considered, as each one associated with $\mu = p_j$ contains p_j as a factor, the product of two or more such divisors, if it divides m, divides one of the d's given above. Consequently a residual polynomial whose coefficient is m divided by this product is expressible as a single variable member of the set S multiplied by a polynomial with integral coefficients.

COROLLARY 2. *For a modulus equal to the power of a prime, a basis is composed of all terms m and* (9).

The divisors of $m = p^k$ are powers of p, and the μ's are all multiples of p.

MOUNT HOLYOKE COLLEGE,
SOUTH HADLEY, MASS.

ON POTENTIALS OF POSITIVE MASS*

PART I

BY

GRIFFITH C. EVANS

I. Introduction

A recent contribution to potential theory is characterized by the names of Lebesgue, Wiener, Kellogg, Vasilesco and Bouligand. Central features of this contribution are the notions of capacity and of regular boundary point, which are related to each other by Kellogg's hypothetical lemma.† A recent memoir by de la Vallée Poussin reinterprets these theories in the light of potentials of positive mass and the Poincaré sweeping out process.‡ In the present memoir, the author's aim is to push on the development of these central problems of mass distribution, regularity, capacity, and approximation, and to answer definitely some of the questions which have become important. Fortunately there is already at hand, in the analysis of the general integral or linear functional of Radon, Daniell, and F. Riesz, the precise mathematical tool which is necessary.§

1. **Integrals and potentials.** Let F be a closed bounded point set, T the infinite domain lying in the complement of F, whose boundary t consists entirely of points of F. We consider an arbitrary distribution of positive mass $f(e)$ on F, finite in total amount.|| The potential of this mass, at a point M, is given by the generalized Stieltjes integral

$$U(M) = \int \frac{1}{MP}\, df(e_P)$$

* Preliminary reports presented to the Society at the meetings of December 29, 1932, October 28, 1933, and December 27, 1933; received by the editors June 25, 1934. The material except §22, Part II, has been delivered as a course of lectures at the Rice Institute in the Spring and Fall of 1933.

† O. D. Kellogg, *Foundations of Potential Theory*, Berlin, 1929, p. 337.

‡ C. de la Vallée Poussin, *Extension de la méthode du balayage de Poincaré, et problème de Dirichlet*, Annales de l'Institut Henri Poincaré, vol. 2 (1932), pp. 169–232.

§ G. Radon, *Theorie und Anwendungen der absolut additiven Mengenfunktionen*, Sitzungsberichte der Akademie der Wissenschaften in Wien, vol. 122, IIa (1913), pp. 1295–1438; *Über die Randwertaufgaben beim logarithmischen Potential*, ibid., vol. 128 (1919), pp. 1123–1167.

P. J. Daniell, (1) *A general form of integral*, Annals of Mathematics, vol. 19 (1918), pp. 279–294; (2) *Further properties of the general integral*, ibid., vol. 21 (1920), pp. 203–220.

F. Riesz, *Über lineare Funktionalgleichungen*, Acta Mathematica, vol. 41 (1916–18), pp. 71–98.

|| That is, $f(e)$ is an additive, bounded, not negative function of Borel measurable point sets e, such that $f(e\cdot CF)=0$, CF being the complement of F.

extended over the whole of space; when necessary we denote this entire space by W, so that $W-F$, $W-t$, etc. have meaning.

Let s be a closed bounded set which is the boundary of a domain Σ; as a whole or in part s may be the boundary also of other domains B_1, B_2, $\cdots$, which thus constitute in their totality an open set B. The set $s+B$ is a closed set G which may or may not have points in common with F.

It is of frequent application that if $f(e)$ is a distribution of positive mass on F, bounded in total amount, the same is true of the function

$$f'(e) = f(e\cdot E)$$

where E is any fixed set, measurable Borel. Also if $f'(e)$, $f''(e)$ are two such functions, with $f(e)=f'(e)+f''(e)$, and $\phi(P)$ a continuous point function,

$$\int_W \phi(P)df(e) = \int_W \phi(P)df_1'(e) + \int_W \phi(P)df''(e).$$

The same equation remains valid for the generalized integral, $\phi(P)$ not being continuous, as far as the integrals exist.

Let now $f_1(e)$, $f_2(e)$, $\cdots$ form a denumerable sequence of such functions chosen so that $f_1(F)+f_2(F)+\cdots$ is convergent. The functions

$$f^n(e) = f_1(e)+f_2(e)+\cdots+f_n(e),\ f(e) = f_1(e)+f_2(e)+\cdots,$$
$$r^n(e) = f_{n+1}(e)+f_{n+2}(e)+\cdots$$

are distributions of positive mass on F, finite in total amount, and if $\phi(P)$ is continuous,

$$\int_W \phi(P)df(e) = \lim_{n=\infty}\int_W \phi(P)df^n(e). \tag{1}$$

In fact,

$$\int_W \phi(P)df(e) = \int_W \phi(P)df^n(e) + \int_W \phi(P)dr^n(e),$$

and, as in (3), below,

$$\left|\int_W \phi(P)dr^n(e)\right| \leqq r^n(F)(\text{u.b. on } F \text{ of } |\phi(P)|).$$

We introduce the function

$$\begin{aligned} h^N(M,P) &= 1/(MP), \qquad & MP \geqq 1/N,\\ &= N, & MP < 1/N. \end{aligned}$$

This is a continuous function of P to which (1) applies. Hence, for all N,

$$\int_W h^N(M, P)df(e_P) = \lim_{n=\infty} \int_W h^N(M, P)df^n(e_P) \leqq \liminf_{n=\infty} \int_W \frac{1}{MP}\, df^n(e_P),$$

and

$$\int_W \frac{1}{MP}\, df(e_P) \leqq \liminf_{n=\infty} \int_W \frac{1}{MP}\, df^n(e_P),$$

admitting $+\infty$ as a possible value of the integral. Also

$$\int_W h^N(M, P)df(e_P) \geqq \int_W h^N(M, P)df^n(e_P) \text{ and}$$

$$\int_W \frac{1}{MP}\, df(e_P) \geqq \int_W \frac{1}{MP}\, df^n(e_P)$$

for every n. Hence

$$\int_W \frac{1}{MP}\, df(e_P) \geqq \limsup_{n=\infty} \int_W \frac{1}{MP}\, df^n(e_P).$$

From the two inequalities we deduce

$$(2) \qquad \int_W \frac{1}{MP}\, df(e_P) = \lim_{n=\infty} \int_W \frac{1}{MP}\, df^n(e_P),$$

which is an identity in the potential functions of the various masses.

Let E be a set measurable Borel. If $\phi(P)$ is *bounded and continuous in* W and $|\phi(P)| \leqq N$ on E, we have

$$(3) \qquad \left| \int_W \phi(P)df(E\cdot e_P) \right| \leqq Nf(E),$$

and if $\phi_1(P)$ and $\phi_2(P)$ are *bounded and continuous in* W *and identical on* E, then

$$(4) \qquad \int_W \phi_1(P)df(E\cdot e_P) = \int_W \phi_2(P)df(E\cdot e_P).$$

So much is seen, for (3), by taking the Stieltjes integral as the limit of a Riemann sum on a net; and (4) is an immediate consequence of (3).

But further, *if the set* E *is merely measurable Borel, and* ϕ, ϕ_1, ϕ_2 *are merely measurable Borel in* W, *then* (3) *is valid if* $|\phi(P)| \leqq N$ *on* E, *and* (4) *if* $\phi_1(P) = \phi_2(P)$ *on* E, *provided the integrals are convergent.*

In fact, if we define the class T_0, of Daniell, as the class of functions $\phi(P)$ on E which can be extended so as to be bounded and continuous in W,

and take the range of the variable P as the set E, the integral defined by the equation

$$I = \int_E \phi(P)df(e) = \int_W \phi(P)df(E\cdot e_P)$$

is uniquely determined by the values of $\phi(P)$ on E, satisfies the postulates (C), (A), (L), (P) of Daniell,* and is therefore a general I-integral on E. It is merely the postulate (L) which requires attention:

(L) If $\phi_1(P) \geqq \phi_2(P) \geqq \cdots$ on E and $\lim\ (n=\infty)\phi_n(P)=0$ for all P in E, then $\lim\ (n=\infty)\ I(\phi_n)=0$.

If E is closed, the $\phi_n(P)$ approaches 0 uniformly on E, and (3), for continuous functions, yields the conclusion (L). If E is not closed, and $\phi_1(P) \leqq N_1$ on E, the lesser, nevertheless, of the two values $\phi_1(P)$, N_1 forms a continuous function $\psi(P)$ on W. Moreover $I(\psi)=I(\phi_1)$ by (4). Hence without loss of generality we may suppose $\phi_n(P) \leqq N_1$ in W. Now, given $\epsilon>0$, E contains a closed set E_1, such that $f(E)-f(E_1)<\epsilon/N_1$; moreover $f(e\cdot[E-E_1])$ is a positive distribution of mass. We have

$$\begin{aligned} I(\phi_n) - I_1(\phi_n) &= \int \phi_n(P)df(E\cdot e_P) - \int \phi_n(P)df(E_1\cdot e_P) \\ &= \int \phi_n(P)df([E - E_1]\cdot e_P) < N_1\epsilon/N_1 = \epsilon. \end{aligned}$$

But $\lim I_1(\phi_n)=0$; whence $\lim I(\phi_n) \leqq \epsilon$. Consequently also $\lim I(\phi_n)=0$.

Let Γ_ρ be a spherical region of radius ρ and center Q, and denote by $U_{\Gamma_\rho}(Q)$ the contribution to the potential at Q due to the mass on Γ_ρ. We note, with de la Vallée Poussin, that *if $U(Q)$ is finite,*

$$\lim_{\rho=0} U_{\Gamma_\rho}(Q) = 0. \tag{5}$$

In fact, given $\epsilon>0$, we can choose N so that

$$U^{(N)}(Q) = \int_W h^N(Q, P)df(e_P) > U(Q) - \epsilon/2.$$

* Daniell, loc. cit. (1), p. 280. The postulates (C), (A), (P) are as follows: (C) If $\theta(P)=c\phi(P)$ on E, $I(\theta)=cI(\phi)$; (A) If $\theta(P)=\phi_1(P)+\phi_2(P)$ on E, $I(\theta)=I(\phi_1)+I(\phi_2)$; (P) If $\phi(P)\geqq 0$ on E, $I(\phi)\geqq 0$. If these postulates are satisfied for functions of the class T_0, which has to be closed with respect to the operations of addition, multiplication by a constant, logical addition and logical multiplication, they remain as properties for the class of functions to which the operation I is generalized. This class in our case contains all those functions on E which arise from functions bounded and measurable Borel on W.

Hence, with this value of N, since the integrands are positive,

$$\int h^N(Q, P)df(\Gamma_\rho \cdot e_P) > U_{\Gamma_\rho}(Q) - \epsilon/2, \text{ for all } \rho.$$

But $h^N(Q, P)$ is bounded and $\lim (\rho=0) f(\Gamma_\rho)=0$, since otherwise $U(Q)$ would be $+\infty$; and we can therefore choose $\rho_0>0$ so that

$$\int h^N(Q, P)df(\Gamma_\rho \cdot e_P) < \epsilon/2, \qquad \rho \leqq \rho_0.$$

Whence

$$U_{\Gamma_\rho}(Q) < \epsilon, \qquad \rho \leqq \rho_0.$$

2. **Potentials and superharmonic functions.** A function $u(M)$ is superharmonic in a bounded domain Ω if (i) it is lower semicontinuous in Ω, not identically equal to $+\infty$, and (ii) given any regular closed surface σ, contained with its interior in Ω, and any function $v(M)$ continuous within and on σ, harmonic within σ, such that $v(M) \leqq u(M)$ on σ, then also $v(M) \leqq u(M)$ within σ.† For (ii) may be substituted the statement that $u(M)$ is at least as great as its mean on any spherical surface of center M and radius ρ, for all ρ sufficiently small, depending on M. Instead of a spherical surface, a spherical volume may be used.

It will be convenient to speak of a function simply as superharmonic if it is superharmonic in every bounded domain, and as superharmonic in an infinite domain if it is superharmonic in every bounded domain contained in the infinite domain.

Let Ω_1 be any domain contained with its boundary Ω_1^* in Ω. A function which is superharmonic in Ω cannot take on its lower bound b for Ω_1 at any point of Ω_1 unless it is identically constant. On the other hand, given $\epsilon>0$, the set of points in $\Omega_1+\Omega_1^*$ where $u(M) \leqq b+\epsilon$ is closed and not empty. Hence $u(M)$ must take on the value b at some point of Ω_1^*. It is F. Riesz's fundamental theorem that $u(M)$ may be written in Ω_1 as the sum of a potential of a distribution of positive mass on Ω_1, finite in total amount, and a function harmonic in Ω_1.‡

The potential $U(M)$ satisfies the conditions (i), (ii), and accordingly is superharmonic. Conversely, *if $u(M)$ satisfies the conditions* (i), (ii) *in W, is harmonic outside F and vanishes continuously at ∞, it must be the potential*

† For the analysis of superharmonic functions, see F. Riesz, (1) *Sur les fonctions subharmoniques et leur rapport à la théorie du potentiel*, Acta Mathematica, vol. 48 (1926), pp. 329–343; (2) *same title*, ibid., vol. 54 (1930), pp. 321–360.

‡ A proof of F. Riesz's theorem is given in §4, below.

$U(M)$ of a distribution of positive mass on F, finite in total amount.†

In order to prove this statement we apply the theorem of F. Riesz to a spherical region Ω_1, of radius ρ and center at a point of F, which contains F in its interior. Since $u(M)$ is harmonic outside F the distribution of positive mass must lie on F. Given $\epsilon>0$, we take ρ so large that the potential $U(M)$ of this distribution is $\leqq\epsilon/2$ on Ω_1^*, and also so large that $|u(M)|$ is $\leqq\epsilon/2$ on Ω_1^*, and write, for all M in W,

$$u(M) = U(M) + v(M).$$

Then $v(M)$ is harmonic in Ω_1 and continuous in $\Omega_1+\Omega_1^*$, and $|v(M)|\leqq\epsilon$ on Ω_1^*. Hence $|v(M)|\leqq\epsilon$ in Ω_1. But this means that $v(M)\equiv 0$.

Lemma. *If $U_1(M)$ and $U_2(M)$ are potentials of positive mass distributions $f_1(e)$, $f_2(e)$ on F, and $U_2(M)\geqq U_1(M)$, for all M, then $f_2(F)\geqq f_1(F)$.*

In fact, if $U_2(M)\equiv U_1(M)$ in T their total masses are equal:

$$\int_C \frac{\partial U_1}{\partial n}\, d\sigma = \int_C \frac{\partial U_2}{\partial n}\, d\sigma,$$

C being a smooth surface sufficiently large to contain F in its interior.‡ If $U_2(M_1)>U_1(M_1)$ for some M_1 in T, it follows that $U_2(M)>U_1(M)$ for all M in T, for $U_2(M)\geqq U_1(M)$ and both functions are harmonic in T. Hence by taking C as a spherical surface of radius ρ sufficiently large and developing U_2-U_1 about ∞, we see that

$$\int_C \frac{\partial U_2}{\partial n}\, d\sigma > \int_C \frac{\partial U_1}{\partial n}\, d\sigma,$$

and

$$f_2(F) > f_1(F).$$

2.1. **Increasing sequences of potentials.** A function which is the limit of an increasing sequence of superharmonic functions and is not identically $+\infty$ is also superharmonic. A similar proposition is the following:

Theorem. *A function $u(M)$ which is the limit of an increasing sequence of potential functions $U_n(M)$ of positive masses $f_n(e)$ on F, such that $f_n(F)$ is bounded, independently of n, is also a potential of a positive distribution $f(e)$ on F, with*

$$f(F) = \lim_{n=\infty} f_n(F).$$

† An obvious generalization of de la Vallée Poussin's theorem, loc. cit., p. 210.

‡ Where n stands for the direction of the normal to a simple closed surface it is taken as positive towards the interior.

In fact, $u(M)$ is not identically $+\infty$ since the convergence of the sequence is uniform outside a sufficiently large sphere. Hence $u(M)$ is superharmonic. Moreover, as is well known,† there exists (by the Cantor "diagonal process") a subsequence $f_{n_i}(e)$ and a rectangular net R, such that for every mesh ω of R we have

$$f(\omega) = \lim_{n_i=\infty} f_{n_i}(\omega), \tag{6}$$

where $f(e)$ is a certain distribution of positive mass on F. And if $\phi(P)$ is any continuous function,

$$\lim_{n_i=\infty} \int_W \phi df_{n_i} = \int_W \phi df. \tag{7}$$

The equations (6), (7) are characteristic of what Radon calls *weak convergence*.‡

Let $U(M)$ be the potential of $f(e)$. We show that $U(M)\equiv u(M)$, by using the mean value on a spherical surface.

2.2. **Mean value on a spherical surface.** Denote by $A_u(\rho, Q)$ the average value of a function $u(M)$ on a spherical surface $C(\rho, Q)$ of radius ρ and center Q:

$$A_u(\rho, Q) = \frac{1}{4\pi\rho^2}\int_{C(\rho,Q)} u(M)dM. \tag{8}$$

We have, for the potential $U(M)$,

$$\begin{aligned} A_U(\rho, Q) &= \frac{1}{4\pi\rho^2}\int_{C(\rho,Q)} dM \int_W \frac{1}{MP}\,df(e_P) \\ &= \frac{1}{4\pi\rho^2}\int_W df(e_P)\int_{C(\rho,Q)} \frac{1}{MP}\,dM, \\ A_U(\rho, Q) &= \int_W h^{1/\rho}(Q, P)df(e_P) \end{aligned} \tag{8'}$$

from the fact that $\int_C [1/(MP)]dM$ is the potential at P of a uniform distribution on $C(\rho, Q)$. But $h^{1/\rho}(Q, P)$ is a continuous function of P.

For any superharmonic function u, we have

$$u(Q) = \lim_{\rho=0} A_u(\rho, Q). \tag{9}$$

For $u(Q) \leqq \lim\inf\ (\rho=0)A_u(\rho, Q)$, being lower semicontinuous, and $u(Q) \geqq A_u(\rho, Q)$ from the property (ii); that is, $u(Q)$ has what may be called the "super-mean" property. Hence $\lim\ (\rho=0)A_u(\rho, Q)$ exists and satisfies (9).

† J. Radon, *Über lineare Funktionalgleichungen*, Sitzungsberichte der Akademie der Wissenschaften in Wien, vol. 128 (1919), pp. 1083–1121, p. 1092.

‡ Ibid., p. 1088. The equation (7) is, in fact, Radon's definition of weak convergence. For the sake of its more evident relation to the structure of the Stieltjes integral, however, we shall say that the sequence of positive mass distributions $\{f_n(e)\}$ converges weakly to $f(e)$ if $f_n(W)$ is bounded, independently of n, and $\lim\ (n=\infty)\, f_n(\omega)=f(\omega)$ for every mesh ω of some rectangular net.

Moreover, from (ii), it is evident that if $u(M)$ is superharmonic, $A_u(\rho, Q)$ is monotone increasing as ρ tends to zero. In particular, $U(Q) = \lim\ (\rho = 0)$ $A_U(\rho, Q)$.

Returning to the statement to be proved, we note that from (8), since the U_n form a monotonic sequence,

$$A_u(\rho, Q) = \lim_{n=\infty} A_{U_n}(\rho, Q),$$

and from (8′), by means of the weak convergence property (7),

$$A_U(\rho, Q) = \lim_{n_i=\infty} A_{U_{n_i}}(\rho, Q).$$

Hence $A_u(\rho, Q) = A_U(\rho, Q)$, and accordingly, by (9),

$$u(Q) = U(Q).$$

But this is what was to be proved. If we take the integral over a large spherical surface C, we have $\int_C (du/dn)d\sigma = \lim(n = \infty)\int_C (dU_n/dn)d\sigma$, or $f(F) = \lim(n = \infty) f_n(F)$.

Evidently a sufficient condition that $f_n(F)$ be bounded is that $u(M)$ be not identically infinite. For in that case $u_n(M)$ will converge uniformly to $u(M)$ in any closed region which has no points in common with F.

3. **Generalized derivatives.** We recall some properties of generalized derivatives, with special relation to the potential function.† Since the space integral of $1/(MP^2)$ extended over a bounded domain Ω is convergent,—in fact, for every P,

$$\int_\Omega \frac{1}{MP^2}\, dM \leqq 4\pi d \tag{10}$$

where d is the diameter of the domain,—it is easily verified that the quantity $U_\alpha(M)$,

$$U_\alpha(M) = \int_W \frac{\cos(MP, \alpha)}{MP^2}\, df(e_P),$$

α being a fixed direction in space, is a summable function, spatially, of M, and that

$$\int_\Omega dM \int_W \frac{\cos(MP, \alpha)}{MP^2}\, df(e_P) = \int_W df(e_P) \int_\Omega \frac{\cos(MP, \alpha)}{MP^2}\, dM. \tag{11}$$

† G. C. Evans, *Complements of potential theory*, II, American Journal of Mathematics, vol. 55 (1933), pp. 29–49.

We denote by $\Phi_\alpha(e)$ the function of point sets generated by $\int_e U_\alpha(M)dM$. *It is absolutely continuous.* Moreover, since wherever $U_x(M)$, $U_y(M)$, $U_z(M)$ exist,

$$(12) \qquad U_\alpha(M) = U_x(M)\cos(x, \alpha) + U_y(M)\cos(y, \alpha) + U_z(M)\cos(z, \alpha),$$

it follows that

$$(13) \qquad \Phi_\alpha(e) = \Phi_x(e)\cos(x, \alpha) + \Phi_y(e)\cos(y, \alpha) + \Phi_z(M)\cos(z, \alpha),$$

for every spatially measurable point set e.

We denote by $D_\alpha U$ *the Lebesgue derivative of the absolutely continuous function of point sets* $\Phi_\alpha(e)$; it exists almost everywhere. In particular, from (13), it exists wherever D_xU, D_yU, D_zU exist, and it satisfies the relations

$$(14) \qquad \begin{aligned} D_\alpha U &= D_xU\cos(x, \alpha) + D_yU\cos(y, \alpha) + D_zU\cos(z, \alpha), \\ D_\alpha U &= U_\alpha, \end{aligned}$$

except possibly on a set of spatial measure zero which is independent of α.

Finally, from the convergence of $\int_\Omega dM \int[1/(MP^2)]df(e_P)$ it follows that $\int[1/(MP^2)]df(e_P)$, $\int[\cos(MP, x)/(MP^2)]df(e_P)$, $\int[1/(MP)]df(e_P)$ represent summable functions of x on almost all lines of direction x. It may be easily verified that on any line l where $\int[1/(MP^2)]df(e_P)$ is summable the total mass must be zero, and that on one of these non-exceptional lines l of direction x, if we select a point A where $U(A)$ is finite,

$$\begin{aligned} \int_A^B dx_M \int_W \frac{\cos(MP, x)}{MP^2} df(e_P) &= \int_{W-l} df(e_P) \int_A^B \frac{\cos(MP, x)}{MP^2} dx_M \\ &= \int_{W-l}\left[\frac{1}{BP} - \frac{1}{AP}\right] df(e_P) = \int_W \left[\frac{1}{BP} - \frac{1}{AP}\right] df(e_P) \\ &= U(B) - U(A), \end{aligned}$$

so that $U(B)$ exists everywhere on l and is absolutely continuous in x.†

That is to say, *on almost all lines with a given direction* α, $U(M)$ *is absolutely continuous as a function of distance on the line, and its partial derivative exists and has almost everywhere on the line the value*

$$(15) \qquad \frac{\partial U}{\partial \alpha} = D_\alpha U = U_\alpha.$$

That this relation holds almost everywhere in space follows from the spatial

† A generalization of these formulas, for integration along a curve, is given by W. H. Binney, *An elliptic system of integral equations on summable functions*, in the present number of these Transactions, pp. 254–265; see Lemma C.

measurability of the partial derivative numbers. The exceptional set, however, is not shown to be independent of α.

For a surface σ, bounding a domain Ω, sufficiently smooth for Green's theorem to hold, we have

$$\int_\Omega D_\alpha U dM = \int_\Omega \frac{\partial U}{\partial \alpha} dM = \int_\Omega U_\alpha dM = \int_\sigma U \cos(n, \alpha) dM, \tag{16}$$

which is a sort of three-dimensional statement of absolute continuity along the direction α. The author has used the phrase "U is a potential of its vector or generalized derivative $D_\alpha U$" to signify the relation

$$\int_\Omega D_\alpha U dM = \int_\sigma U \cos(x, \alpha) dM, \text{ for all } \alpha, \tag{16'}$$

even if $\partial U / \partial \alpha$ may fail to exist.

By Riesz's fundamental theorem, it follows that (16) applies to a function u for a bounded region Ω contained with its smooth boundary within any domain in which u is superharmonic.

4. **Further properties of the average. Riesz's theorem.** In order to illustrate further useful properties of the average we shall give a brief proof of the theorem on the resolution of a superharmonic function into a potential and a harmonic function, in line with the ideas of F. Riesz and T. Radó.

Denote by $u(\rho, Q)$ or $a_u(\rho, Q)$ or, for brevity, u_ρ, the average of a function $u(M)$ over a spherical region $\Gamma(\rho, Q)$ of center Q and radius ρ; we take the same ρ for all Q. Similarly, denote by $u(\rho', \rho'', Q)$ or $u_{\rho'\rho''}$ the iterated average, obtained by averaging $u(\rho', M)$ over a sphere of radius ρ'' and center Q. We have

$$u(\rho, Q) = \frac{3}{\rho^3} \int_0^\rho A_u(\rho, Q) \rho^2 d\rho, \tag{17}$$

A_u being the mean on the spherical surface, as before. Significant properties of these averages are well known and easily established.† If $u(M)$ is summable in a domain Ω, $u(\rho, M)$ is continuous in any portion of Ω distant from the boundary of Ω by as much as ρ, and $\lim (\rho = 0)\ u(\rho, M) = u(M)$ for almost all M; $\partial u(\rho, M)/\partial x$, $\partial u(\rho, M)/\partial y$, $\partial u(\rho, M)/\partial z$ exist and equal the generalized derivatives almost everywhere; if $u(M)$ is a potential function of its generalized or vector derivative, as in (16′), §3, then

† H. E. Bray, *Proof of a formula for an area*, Bulletin of the American Mathematical Society, vol. 29 (1923), pp. 264–270; *Green's lemma*, Annals of Mathematics, vol. 26 (1925), pp. 278–286; T. Radó, *Remarques sur les fonctions subharmoniques*, Comptes Rendus de l'Académie des Sciences, vol. 186 (1928), pp. 346–348; F. Riesz, Memoir (2) cited in §2, see p. 342 ff., where other references are given.

$$a_{D_\alpha u}(\rho, Q) = D_\alpha a_u(\rho, Q) = \frac{\partial u(\rho, Q)}{\partial \alpha}, \tag{18}$$

these quantities being continuous. Hence by iterating the averaging operation the derivatives to any order may be made continuous.

If $u(M)$ is superharmonic in Ω, $A_u(\rho, M)$ and $a_u(\rho, M)$ are superharmonic in any portion of Ω distant from the boundary of Ω by more than ρ. In fact, if we write $u(Q)=u(0, 0, 0)$, $u(M)=u(x, y, z)$, $u(P)=u(x+\xi, y+\eta, z+\zeta)$, and formulate explicitly the averages, we see that the successive averaging operations are commutative; that is, for all ρ' sufficiently small

$$\begin{aligned} u(\rho, \rho', Q) &= \frac{3}{4\pi\rho'^3}\int_{\Gamma(\rho',0)} dx\,dy\,dz \frac{3}{4\pi\rho^3}\int_{\Gamma(\rho,0)} u(x+\xi, y+\eta, z+\zeta)d\xi\,d\eta\,d\zeta \\ &= u(\rho', \rho, Q). \end{aligned} \tag{19}$$

Now $u(\rho, M)$ is continuous in M; moreover from (ii), §2,

$$u(\rho, Q) = \frac{3}{4\pi\rho^3}\int_{\Gamma(\rho,Q)} u(M)dM \geqq \frac{3}{4\pi\rho^3}\int_{\Gamma(\rho,Q)} u(\rho', M)dM = u(\rho', \rho, Q),$$

so that, from (19),

$$u(\rho, Q) \geqq u(\rho, \rho', Q) = \frac{3}{4\pi\rho'^3}\int_{\Gamma(\rho',Q)} u(\rho, R)dR$$

which is a substitute for the condition (ii), and makes $u(\rho, M)$ superharmonic. A similar demonstration applies to the A-operation.

Again let $u(M)$ be superharmonic in Ω. The function $u(\rho_1, \rho_2, \cdots, \rho_k, Q)$ is a weighted mean of $u(M)$ over a sphere $\Gamma(\rho_1+\rho_2+\cdots+\rho_k, Q)$, such that $u(\rho_1, \rho_2, \cdots, \rho_k, Q) \leqq u(Q)$. Moreover, since $u(M)$ is lower semicontinuous, $u(Q) \leqq \liminf\,(M=Q)u(M)$; hence

$$\lim u(\rho_1, \rho_2, \cdots, \rho_k, Q) = u(Q),\ Q \text{ in } \Omega,$$

as $\rho_1, \rho_2, \cdots, \rho_k$ tend independently to zero. In particular,

$$\lim_{\rho=0} u^{(k)}(\rho, Q) = u(Q),\ Q \text{ in } \Omega, \tag{19'}$$

where $u^{(k)}(\rho, M)=u(\rho_1, \rho_2, \cdots, \rho_k, M)$ with $\rho_1=\rho_2=\cdots=\rho_k=\rho$, and increases monotonically as ρ decreases to zero.

Let D be a bounded domain, Ω a domain contained with its boundary Ω^* in D. Let Ω_i, $i=1, 2, 3$, be intermediate domains with boundaries Ω_i^* such that we have

$$\Omega + \Omega^* \text{ in } \Omega_1;\ \Omega_i + \Omega_i^* \text{ in } \Omega_{i+1},\ i = 1, 2;\ \Omega_3 + \Omega_3^* \text{ in } D,$$

and Ω_i^* is sufficiently smooth for application of Green's Theorem to regions bounded internally or externally by it.

THEOREM OF F. RIESZ. *If $u(M)$ is superharmonic in D, it may be written in the form*

$$u(M) = U(M) + v(M), \quad M \text{ in } \Omega,$$

where $U(M)$ is the potential of a distribution of positive mass on Ω, finite in total amount, and $v(M)$ is harmonic in Ω.

Let $u^{(1)}(M), u^{(2)}(M), \cdots, u^{(p)}(M), \cdots$ be a monotonic-increasing sequence of continuous superharmonic functions, with limit $u(M)$, in a region Ω_4 which contains $\Omega_3+\Omega_3^*$; for instance, let $u^{(p)}(M)$ be the average $u(1/(p_0+p), M)$, with p_0 fixed and sufficiently great. Define

$$\begin{aligned} u_p(M) &= u^{(p)}(M), \quad M \text{ in } \Omega_1 + \Omega_1^*, \\ &= w_p(M), \quad M \text{ in } \Omega_3 - (\Omega_1 + \Omega_1^*), \\ &= u^{(p)}(M), \quad M \text{ in } \Omega_4 - \Omega_3, \end{aligned}$$

where $w_p(M)$ is the function which is harmonic in $\Omega_3-(\Omega_1+\Omega_1^*)$ and takes on continuously the boundary values $u^{(p)}(M)$ on Ω_1^* and Ω_3^*.

Then $u_p(M)$ is continuous in Ω_4 and evidently possesses the super-mean property. It is therefore superharmonic in Ω_4. Moreover the sequence $u_p(M)$ is monotonic-increasing, and $\leqq u(M)$, converging accordingly to a function $w(M)$, superharmonic in Ω_4. We note that $w(M)$ is harmonic in $\Omega_3-(\Omega_1+\Omega_1^*)$ and identical with $u(M)$ in $\Omega_1+\Omega_1^*$.

We take 4ρ small in comparison with the distances between Ω_3^* and the boundary of Ω_4, between Ω_3^* and Ω_2^*, and between Ω_2^* and Ω_1^*; and consider the function $w(\rho, \rho, \rho, \rho, M) = w^{(4)}(\rho, M)$, which is superharmonic in a region which contains $\Omega_3+\Omega_3^*$ and is identical with $w(M)$ in a neighborhood of Ω_2^*, where it is harmonic. The function $w^{(4)}(\rho, M)$ has continuous third-order partial derivatives, and tends, increasing monotonically, to $w(M)$ in a domain which includes Ω_3, as ρ tends to zero. By means of Green's theorem, as applied to Ω_2, we have the decomposition

$$w^{(4)}(\rho, M) = U_4(M) + v_0(M), \quad M \text{ in } \Omega_2,$$

where $U_4(M)$ is the potential of a distribution of positive continuous density on Ω_2, finite in total amount, and where $v_0(M)$ is harmonic in Ω_2, continuous on $\Omega_2+\Omega_2^*$. Since $v_0(M)$ involves merely the boundary values of $w^{(4)}$ and $dw^{(4)}/dn$ on Ω_2, it follows that $v_0(M)$ will remain independent of ρ, as ρ tends to zero.

The potential $U_4(M)$ therefore increases monotonically as ρ tends to zero. Moreover, since $U_4(M)$ is identical with $w(M)-v_0(M)$ in the neighbor-

hood of Ω_2, and is harmonic there, the total mass is given by 4π times the integral of the normal derivative of $w(M)-v_0(M)$, and does not involve ρ. Hence by the theorem of §2.1, the function

$$U_1(M) = \lim_{\rho=0} U_4(M)$$

is the potential of a distribution of positive mass $f(e)$ on a bounded set, and is bounded in total amount.

We have

$$u(M) = w(M) = v_0(M) + \int_W \frac{1}{MP}\, df(e_P), \qquad M \text{ in } \Omega.$$

Writing

$$f(e) = f(e\cdot\Omega) + f(e\cdot(W - \Omega)),$$

$$U(M) = \int_W \frac{1}{MP}\, df(e_P\cdot\Omega), \qquad v(M) = v_0(M) + \int_W \frac{1}{MP}\, df(e_P\cdot(W - \Omega)),$$

we shall have

$$u(M) = U(M) + v(M), \qquad M \text{ in } \Omega,$$

according to the conditions of the theorem.

II. Limiting values of potential function

In general terms it seems safe to say that the potential of a positive distribution of mass is greater where mass is than where it is not. In fact, the potential cannot take on its upper bound at a point of positive distance from the mass, being harmonic at such a point. Nevertheless it is evident that a potential need not take on its upper bound at any point whatever in space, and if our naive idea is to be made precise, it must involve the limiting values of potential functions as we approach points of the mass distribution.

We note that for any point M whatever, on account of the lower semicontinuity of $U(M)$, and equation (9), §2.1,

$$U(M) = \liminf_{M'=M} U(M'), \qquad M' \text{ in } W. \tag{1}$$

5. **Points of continuity.** In terms of the notation of §1 we have the following

THEOREM. *Let Q be a point of t, not an isolated point, and $U(Q)$ finite; if*

$$\lim_{P=Q} U(P) = U(Q), \qquad P \textit{ in } t, \tag{2}$$

then

$$\lim_{M=Q} U(M) = U(Q), \qquad M \textit{ in } T. \tag{3}$$

For this theorem, T may be any domain containing no points of F, whose boundary t consists of points of F.

On account of (1) it is sufficient to prove that

$$\limsup_{M=Q} U(M) \leqq U(Q), \qquad M \text{ in } T.$$

We assume that the theorem is false, and that therefore there is a sequence of points $M_1, M_2, \cdots$ in T, and a number $\epsilon > 0$, such that

$$U(M_i) > U(Q) + \epsilon, \qquad \lim_{i=\infty} M_i = Q. \tag{4}$$

Let $\epsilon_1, \epsilon_2, \cdots$ be a sequence of positive numbers such that

$$\epsilon_1 + \epsilon_2 + \cdots < \epsilon/2.$$

Let $\Gamma(\rho, Q)$ be an open spherical neighborhood with center Q and radius ρ, small enough so that the contribution $U_{\Gamma_\rho}(Q)$ to the potential at Q from the mass on $F\cdot\Gamma(\rho, Q)$ satisfies the inequality (see (5) §1)

$$U_{\Gamma_\rho}(Q) < \epsilon_1.$$

It follows therefore that

$$U_{W-\Gamma_\rho}(Q) > U(Q) - \epsilon_1.$$

We suppose also that ρ is small enough so that, for P in $t\cdot\Gamma(\rho, Q)$,

$$U(P) < U(Q) + \epsilon_2.$$

But $U_{W-\Gamma_\rho}(P)$ is continuous at Q for any method of approach; let therefore $\Gamma(\delta, Q)$ be a second spherical neighborhood, concentric with $\Gamma(\rho, Q)$, with $\delta < \rho$, δ being small enough so that, for P in $\Gamma(\delta, Q)$,

$$U_{W-\Gamma_\rho}(P) > U_{W-\Gamma_\rho}(Q) - \epsilon_3 > U(Q) - \epsilon_1 - \epsilon_3.$$

Thus we have

$$\begin{aligned} &U_{\Gamma_\rho}(P) < U(Q) + \epsilon_2 - (U(Q) - \epsilon_1 - \epsilon_3), \\ &U_{\Gamma_\rho}(P) < \epsilon_1 + \epsilon_2 + \epsilon_3, \qquad P \text{ in } t\cdot\Gamma(\delta, Q). \end{aligned} \tag{5}$$

Denote the distance M_iQ by δ_i; we may assume without loss of generality that $\delta_i < \delta/2$ for all i. Let Q_i be a point in F at the minimum distance, say δ_i', from M_i. Such a point Q_i exists, since F is closed; moreover Q_i lies in t, since otherwise the segment Q_iM_i would contain a point of t nearer to M_i

than Q_i. Further, Q_i lies in $\Gamma(\delta, Q)$, since $QQ_i \leqq QM_i + M_iQ_i < \delta/2 + \delta/2$. Finally, $\lim (i=\infty) Q_i = Q$.

We have

$$Q_iP \leqq Q_iM_i + M_iP \leqq 2M_iP, \qquad P \text{ in } F\cdot\Gamma(\rho, Q). \tag{6}$$

In fact, $Q_iM_i = \delta_i'$, $M_iP \geqq \delta_i'$. Also, for P in $F - F\cdot\Gamma(\rho, Q)$,

$$\begin{aligned} Q_iP \leqq Q_iM_i + M_iP &= \left(1 + \frac{Q_iM_i}{M_iP}\right) M_iP \\ &\leqq \left(1 + \frac{\delta_i'}{\rho - \delta_i}\right) M_iP, \end{aligned}$$

since for such P, $M_iP \geqq QP - QM_i \geqq \rho - \delta_i$. Accordingly, since $\rho > \delta > \delta_i \geqq \delta_i'$,

$$Q_iP < \left(1 + \frac{\delta_i}{\rho - \delta}\right) M_iP, \qquad P \text{ in } F - F\cdot\Gamma(\rho, Q). \tag{7}$$

Now, from (6), (7), making use of the properties (C), (P) of the general integral,

$$\begin{aligned} U(M_i) = U_{\Gamma_\rho}(M_i) + U_{W-\Gamma_\rho}(M_i) &= \int_W \frac{1}{M_iP}\, df(e_P\cdot[F\cdot\Gamma(\rho, Q)]) \\ &\quad + \int_W \frac{1}{M_iP}\, df(e_P\cdot[F - F\cdot\Gamma(\rho, Q)]) \\ &< 2U_{\Gamma_\rho}(Q_i) + \left(1 + \frac{\delta_i}{\rho - \delta}\right) U_{W-\Gamma_\rho}(Q_i) \\ &= U_{\Gamma_\rho}(Q_i) + U(Q_i) + \frac{\delta_i}{\rho - \delta}\, U_{W-\Gamma_\rho}(Q_i) \\ &< \epsilon_1 + \epsilon_2 + \epsilon_3 + U(Q) + \epsilon_2 + \frac{\delta_i}{\rho - \delta}\,(U(Q) + \epsilon_2). \end{aligned}$$

Choose now i great enough so that

$$\frac{\delta_i}{\rho - \delta}\,(U(Q) + \epsilon_2) < \epsilon_4.$$

Then

$$U(M_i) < U(Q) + \epsilon_1 + 2\epsilon_2 + \epsilon_3 + \epsilon_4 < U(Q) + \epsilon.$$

But this statement contradicts (4). Accordingly the proof is complete.

COROLLARY. *Let Q be a point of t, isolated or not, and $U(Q)=+\infty$. Then*

$$\lim_{M=Q} U(M) = U(Q), \qquad M \text{ in } W.$$

In fact, this is equation (1) in this case.

5.1. **Frequency of points of continuity.** Let the set t be without isolated points; it is then perfect. Let Γ_1 be any closed spherical region which contains F in its interior. The functions

$$U^{(N)}(P) = \int_W h^N(P, P')df(e_{P'})$$

are, on Γ_1, positive, continuous, and bounded away from zero by a lower bound independent of N, as N tends to $+\infty$. Accordingly the functions

$$v_N(P) = 1/U^{(N)}(P)$$

are continuous and uniformly bounded on Γ_1, and the function

$$v(P) = 1/U(P)$$

is a function of the *first class of Baire* on Γ_1. It is therefore *punctually discontinuous* on Γ_1.† Given any perfect subset E of t (for instance, all the points of t in a closed spherical neighborhood of a point P of t) there will be points of E at which $v(P)$ is continuous, considered only on E. In other words, admitting $+\infty$ as a possible value of U, we have the following proposition:

THEOREM. *If t is perfect, and P is given in t, then in any neighborhood of P there is a point Q of t such that*

$$\lim_{P'=Q} U(P') = U(Q), \qquad P' \text{ in } t.$$

By an application of the theorem of §5 we have the following corollary:

COROLLARY I. *If t is perfect, and P is given in t, then in any neighborhood of P there is a point Q of t such that*

$$\lim_{M=Q} U(M) = U(Q), \qquad M \text{ in } T + t.$$

The following corollary may also be stated, and proved in the same manner as the above. In fact, in the theorem of §5 we may replace t by F, and T by CF, although CF is not necessarily a domain, Q being a frontier point of CF.

† Lebesgue, *Leçons sur l'Intégration*, Paris, 1928, p. 203.

Corollary II. *If F is perfect and P is a point of F, there will be, in any neighborhood of P, a point Q of F such that*

$$\lim_{M=Q} U(M) = U(Q), \qquad M \text{ in } W. \tag{8}$$

This corollary affords an immediate proof of Kellogg's Lemma (see §18).

6. **The superior limit of the potential function.** We return to the set t as in §5, which is closed but not necessarily perfect, and let T be a bounded or an unbounded domain.

Theorem.† *Let Q be a point of t, not an isolated point, and let k, assumed to be finite, be the superior limit of $U(P)$, P in t, as P tends to Q. Given $\epsilon>0$, we can find Q_1 in t, arbitrarily close to Q, so that*

$$\limsup_{M=Q_1} U(M) < U(Q_1) + \epsilon < k + 2\epsilon, \qquad M \text{ in } T. \tag{9}$$

With $\epsilon_1, \epsilon_2, \cdots$ as before, we choose a neighborhood Ω of Q of diameter small enough so that $U(P)<k+\epsilon_2$ for P in $t\cdot\Omega$. For Q_1 we take any point in $t\cdot\Omega$ for which $U(Q_1)>k-\epsilon_5$; this choice of Q_1 is possible by definition of k. The proof of the theorem from this point on is substantially that of the theorem of §5, with Q of §5 replaced by Q_1. Hence it need not be repeated here.

The following simple remarks supplement the theorem just given.

I. *If Q is an isolated point of t, either $U(M)$ is continuous (and harmonic) at Q, or else* $\lim\,(M=Q)U(M)=U(Q)=+\infty$, *for M in W.*

If there is no mass on Q, $U(M)$ is bounded and harmonic in the neighborhood of Q and continuous at Q, therefore harmonic at Q. If there is a point mass at Q, $U(Q)=+\infty$, $\lim\,(M=Q)\,U(M)=+\infty$.

II. *Let Σ be any domain and s its boundary, Q any point of s. Then*

$$\limsup_{P=Q} U(P) \leqq \limsup_{M=Q} U(M), \qquad P, Q \text{ in } s,\ M \text{ in } \Sigma. \tag{10}$$

For there exists a sequence of points P_i of s, tending to Q, such that $\lim U(P_i)$ exists and equals $\limsup\,(P=Q)U(P)$. But there is a sequence of points M_i in Σ, $M_iP_i<\epsilon/2^i$, $U(M_i)>U(P_i)-\epsilon/2^i$, ϵ given >0; for $U(M)$ is lower semicontinuous. Hence, admitting the value $+\infty$ as a possible limit, the equation (10) is established.

† The theorem is slightly sharper than the lemma of §2 of G. C. Evans, *Application of Poincaré's sweeping-out process*, Proceedings of the National Academy of Sciences, vol. 19 (1933), pp. 457–461, but the proof is quite similar. In the cited lemma t was assumed to contain all the mass, instead of being merely a frontier of the mass, as here.

The theorem of §4 is a corollary of the above, by taking $Q_1=Q$.

III. We leave unanswered the question†

$$\text{(10')} \qquad \text{u.b. } U(P)(P \text{ in } t) = \text{u.b. } U(M)(M \text{ in } T)?$$

7. **Inferior limit of the potential function.** We shall see in §7.2 that there may exist points of the mass where the potential is actually less than its lower limit for approach not on the mass. For their consideration there is no gain in restricting the sets in which they lie to points of the mass, specifically, and accordingly we discuss them with reference to the sets s, Σ, B of §1. We are thus dealing with the properties of potentials as positive superharmonic functions rather than as explicitly given integrals.

If there exists a potential $U(M)$ of positive mass on F such that

$$\text{(11)} \qquad U(Q) < \liminf_{M=Q} U(M), \qquad Q \text{ in } s, \ M \text{ in } \Sigma,$$

we say that Q is an *exceptional point of s with respect to* Σ; similarly we speak of exceptional points of s with respect to $\Sigma+B$, etc. As is evident from the definition, we are dealing with geometrical properties of the sets in question.

We use the symbol $C(\rho, Q, E)$ for the portion of the spherical surface $C(\rho, Q)$ which is common to it and a set E, measurable Borel. Such a portion is also measurable Borel.

7.1. **Special cases.** We prove the following theorem:

THEOREM I. *If Q is an exceptional point of s with respect to* Σ,

$$\lim_{\rho=0} \frac{C(\rho, Q, \Sigma)}{C(\rho, Q, s + B)} = 0;$$

also, if Q is an exceptional point of s with respect to $\Sigma+B$,

$$\lim_{\rho=0} \frac{C(\rho, Q. \Sigma + B)}{C(\rho, Q, s)} = 0.$$

Suppose that

$$\liminf_{M=Q} U(M) = U(Q) + h, \qquad M \text{ in } \Sigma, \ h > 0.$$

Let θ be any fixed number $0<\theta<1$. Since $U(M)$ is lower semicontinuous, it follows that, given $\epsilon>0$, there is a spherical neighborhood $\Gamma(\rho_1, Q)$, such that

$$U(M) > U(Q) + \theta h, \qquad M \text{ in } \Sigma, \ QM < \rho_1,$$

$$U(P) > U(Q) - \epsilon, \qquad P \text{ in } W, \ QP < \rho_1.$$

Hence

† The question is now answered in the affirmative, by A. J. Maria, *The potential of a positive mass and the weight function of Wiener*, Proceedings of the National Academy of Sciences, vol. 20 (1934), pp. 485–489.

$$C(\rho, Q)U(Q) > C(\rho, Q, \Sigma)[U(Q) + \theta h] + C(\rho, Q, s + B)[U(Q) - \epsilon],$$
$$0 > \theta h C(\rho, Q, \Sigma) - \epsilon C(\rho, Q, s + B),$$

$$\frac{C(\rho, Q\,\Sigma)}{C(\rho, Q, s + B)} < \frac{\epsilon}{\theta h}, \qquad \rho < \rho_1. \tag{12}$$

Similarly, if

$$\liminf_{M=Q} U(M) = U(Q) + h, \qquad M \text{ in } \Sigma + B,\ h > 0,$$

there exists ρ_2 such that

$$\frac{C(\rho, Q, \Sigma + B)}{C(\rho, Q, s)} < \frac{\epsilon}{\theta h}, \qquad \rho < \rho_2. \tag{13}$$

From this the conclusions of the theorem are evident.

COROLLARY. *An exceptional point of s with respect to Σ is a point of s of spatial density unity in $s+B$. If s is a set of Lebesgue spatial measure zero it has no exceptional points with respect to $\Sigma+B$.*

We remark that the exceptional points of s with respect to Σ must all be regular boundary points of Σ, since an irregular boundary point of Σ cannot be a point of spatial density unity on $s+B$ (see §23, below).

Perfect totally disconnected sets, of which the typical example is the spatial discontinuum, need not be of zero spatial measure. But they are closed sets of *dimension zero* in the sense of Menger. The following theorem about such sets may be proved very simply.

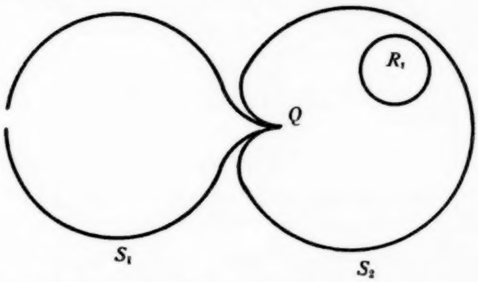
THEOREM II. *If s is of dimension zero at Q, Q cannot be an exceptional point with respect to Σ. If s is of dimension zero, none of its points are exceptional with respect to Σ.*

If s is of dimension zero at Q, in any neighborhood about Q there is contained a neighborhood Ω, to which also Q belongs, whose boundary Ω^* contains no points of s. The set Ω^* is at a positive distance from the closed set s, and, for Ω small enough, since it contains at least one point of Σ, lies entirely in Σ. The lower bound of $U(M)$ for Ω is taken on at some point M_1 of Ω^*. Hence by taking a sequence of neighborhoods with diameters which approach zero as a limit we obtain a sequence of points $M_1, M_2, \cdots$, $\lim M_i = Q$, M_i in Σ, such that $U(Q) \geqq U(M_i)$.

The second part of the theorem is a consequence of the first, for a set of dimension zero is of dimension zero at every one of its points.

7.2. **An exceptional case.** The accompanying figure illustrates a point Q of s for which lim inf $(M=Q)$ $U(M) > U(Q)$, M in Σ. The set B is vacuous, and $F \equiv s$, mass being deposited only on s.

We construct a surface of revolution S_1, made in part of a sphere with center O and in part of an exterior Lebesgue spine with vertex at Q, the sphere being pierced opposite Q on QO, so that the complement of S_1 becomes a domain. To S_1 is applied a conductor potential distribution (see §15) with



potential $U_1(M)$. The point Q being irregular for the domain $W-S_1$, we know, from the symmetry of the figure about OQ, that, as M approaches Q along the extension of OQ (from the right, in the figure),

$$\lim_{M=Q} U_1(M) = U_1(Q) = 1 - k < 1.$$

We now adjoin to the end of the spine, on the same axis of revolution, a closed surface of revolution S_2, exterior to S_1, such that on S_2, in the neighborhood of Q, $U_1(M) \geqq 1-k/2$. Throughout B_2, the region interior to S_2, we distribute a uniform mass, of density sufficiently small so that its potential nowhere exceeds the value $k/8$. Accordingly the inferior limit of the potential of the total mass, for approach to Q, is still obtained by a path through the region B_2. In the figure so far constructed we have $F = B_2+S_1+S_2$.

We may now eliminate the interior mass and still retain the desired property. Denote by σ the projection of B_2 on the x, y plane, and let the points R_k of σ, for which x and y are both rational, be put in countable order. With R_k as center of base, construct a cylindrical surface σ_k of length l, long in comparison with the diameter of S_2, remove the mass from the interior of the tube, and place a uniform mass on the surface of the tube, so that

(i) the potential of this mass at any point in the set O_k common to the interior of the tube σ_k and to B_2 shall be $\geqq 2$,

(ii) the potential of this mass at Q shall be $< 2^{-i}(k/8)$,

(iii) the tube σ_k shall be exterior to the tubes $\sigma_1, \sigma_2, \cdots, \sigma_{k-1}$, if R_k is exterior to them; otherwise R_k shall be omitted from the sequence,

(iv) the sum of the masses on all the σ_k shall be finite. This construction is possible, since a segment of a straight line is a point set of zero capacity in three dimensions; that is, any positive charge on the line will make the potential infinite at some point.

The set $S_1+S_2+B-\Sigma O_k+F'$, where F' denotes the closed cover of $\Sigma\sigma_k$, is closed and bounded. We take this set as the set s, and denote by $U(M)$ the potential of the masses distributed on s. For Σ we have the entire set complementary to s. We have evidently

$$U(Q) \leqq 1 - 3k/4,$$

$$\liminf_{M=Q} U(M) > 1 - k/2, \qquad M \text{ in } \Sigma.$$

Hence

$$\liminf_{M=Q} U(M) > U(Q) + k/4, \qquad M \text{ in } \Sigma.$$

Incidentally, we know that s is of positive spatial measure, of set density 1 at Q, and that Q is a regular boundary point of Σ.

7.3. A theorem on the inferior limit. We prove the following

THEOREM. *Let Q be a point of s, x, y two arbitrary directions at right angles, Ω an arbitrary neighborhood of Q. Then Q is not an exceptional point of s with respect to $\Sigma+B$ unless the set $\Omega\cdot s$ contains a family of closed rectangular contours, with directions parallel to x, y, whose vertices constitute a set of positive spatial measure.*

Let z be a direction normal to x, y. Denote by $\Gamma(\rho, Q, E)$ the portion of $\Gamma(\rho, Q)$ which lies in the Borel measurable set E. Then

$$\Gamma(\rho, Q, E) = \int_0^\rho C(\rho, Q, E)d\rho. \tag{14}$$

Suppose the theorem is not true, and thus that

$$\liminf_{M=Q} U(M) = U(Q) + 3h, \qquad M \text{ in } \Sigma + B,\ h > 0. \tag{15}$$

Let E denote the set of points in W where $U(M) \leqq U(Q)+h$; it is closed.

Given $\eta>0$, there exists $\rho_0>0$ so that for $\rho \leqq \rho_0$, $\Gamma(\rho, Q, E)$ is contained in Ω and

$$\begin{aligned} &\text{meas } \Gamma(\rho, Q, E) \geqq (1-\eta)4\pi\rho^3/3,\\ &\Gamma(\rho, Q, E) \text{ lies in } s,\\ &\liminf_{M=P} U(M) \geqq U(P) + h, \qquad M \text{ in } \Sigma + B,\ P \text{ in } \Gamma(\rho, Q, E). \end{aligned} \tag{16}$$

In fact, by (15), for ρ sufficiently small, $U(M)>U(Q)+2h$, M in $\Sigma+B$, $QM<\rho$. Hence $\Gamma(\rho, Q, E)$ lies in s, and the second and third of conditions (16) are satisfied, for such ρ. Now, similarly to (13), for ρ sufficiently small, given $\epsilon>0$,

$$C(\rho, Q, W-E) < \frac{\epsilon}{h} C(\rho, Q, E),$$

$$C(\rho, Q) < \frac{\epsilon+h}{h} C(\rho, Q, E),$$

$$\Gamma(\rho, Q) < \frac{\epsilon+h}{h} \Gamma(\rho, Q, E),$$

from which follows the first of equations (16).

On almost any plane $z=\text{const.}$, $U(M)$ is continuous along almost all lines $x=\text{const.}$ and $y=\text{const.}$ (see §3). Select then a non-exceptional plane $z=z_0$, which cuts $\Gamma(\rho_0, Q, E)$ in a set of positive superficial measure, and in this plane a line $y=y_0$, which cuts $\Gamma(\rho_0, Q, E)$ in a set E_x of positive linear measure. On almost all lines $x=\text{const.}$, which pass through points of E_x, $U(M)$ is continuous.

There is a closed subset F_x of E_x of positive measure, such that $\lim\,(y=y_0)\; U(x, y, z_0)=U(x, y_0, z_0)$ uniformly, for x in F_x; that is, $\delta'>0$ exists so that

$$(17)\qquad U(x, y, z_0) - U(x, y_0, z_0) < h/4, \qquad |y-y_0| \leqq \delta', \qquad x \text{ in } F_x.$$

In fact, let $\{y_i\}$ be a sequence of values, tending to y_0. Since $U(x, y, z_0)$ is lower semicontinuous, the sets of points in a rectangle $a<x<b$, $y_0 \lesssim y \lesssim y_i$, contained in $\Gamma(\rho, Q)$, where $U(x, y, z_0)>c$ and where $U(x, y, z_0)<c$ respectively, for any c, are measurable Borel; hence their projections on $y=y_0$ are also measurable Borel. But these are respectively the sets where $f_i(x)>c$ and $\phi_i(x)<c$, $f_i(x)$ being the upper bound and $\phi_i(x)$ the lower bound of $U(x, y, z_0)$ considered as a function of y, $y_0 \lesssim y \lesssim y_i$, for x in the interval $a<x<b$. Hence $f_i(x)$ and $\phi_i(x)$ are measurable Borel, and

$$\lim_{i\to\infty} f_i(x) = U(x, y_0, z_0), \qquad \lim_{i\to\infty} \phi_i(x) = U(x, y_0, z_0), \qquad x \text{ in } E_x.$$

But, by Egoroff's theorem, corresponding to the sequence $\{i\}$, there is a closed subset F_x of E_x, of measure differing arbitrarily little from that of E_x, such that the approach of $f_i(x)$ and $\phi_i(x)$ to their limiting values is uniform, for x in F_x. Also, for $y_0 \lesssim y \lesssim y_i$,

$$\phi_i(x) - U(x, y_0, z_0) \leqq U(x, y, z_0) - U(x, y_0, z_0) \leqq f_i(x) - U(x, y_0, z_0),$$

so that the approach of the middle member of this inequality to zero is uniform for x in F_x.

For almost all the points of F_x the linear set density of F_x is unity. Select one of these non-exceptional points (x_0, y_0, z_0). For almost all lines $y=$const. in the plane $z=z_0$, $U(M)$ is continuous. We may therefore apply the sort of argument just used in the case of F_x to the neighborhood of y_0 on the line $x=x_0$. There exists then a set F_y in y, $|y-y_0|\leqq\delta'$, of positive measure on the line $x=x_0$, and a $\delta''>0$ such that

$$(18)\quad U(x, y, z_0) - U(x_0, y, z_0) < h/4, \qquad \text{for } |x - x_0| < \delta'', \qquad y \text{ in } F_y.$$

We may choose the (x_0, y_0, z_0), δ', δ'' so that the entire figure lies in the sphere of center Q and radius ρ_0.

Consider then a rectangle in $z=z_0$, composed of two lines $x=x_1$, $x=x_2$, x_1, x_2 in F_x and two lines $y=y_1$, $y=y_2$, y_1, y_2 in F_y, with $|x_1-x_0|$, $|x_2-x_0|<\delta''$, and $|y_1-y_0|$, $|y_2-y_0|<\delta'$. This rectangular contour lies entirely in s. In fact, for any point on a side parallel to x, say (x, y_1, z_0), we have

$$\begin{aligned} U(x, y_1, z_0) &- U(x_0, y_0, z_0) \\ &\leqq U(x, y_1, z_0) - U(x_0, y_1, z_0) + U(x_0, y_1, z_0) - U(x_0, y_0, z_0) \\ &< h/4 + h/4 = h/2, \end{aligned}$$

by (17), (18). But by the third of equations (16), since $P=(x_0, y_0, z_0)$ lies in $\Gamma(\rho_0, Q, E)$, $M=(x, y_1, z_0)$ cannot lie in $\Sigma+B$. Also for any point on a side parallel to y, say (x_1, y, z_0), we have

$$U(x_1, y, z_0) - U(x_1, y_0, z_0) < h/4.$$

But (x_1, y_0, z_0) lies in $\Gamma(\rho_0, Q, E)$; hence (x_1, y, z_0) cannot lie in $\Sigma+B$.

The set of points (x_1, y_1, z), (x_2, y_2, z), (x_1, y_2, z), (x_2, y_1, z), vertices of rectangular contours which lie in s, for which $|x_2-x_1|\geqq a$, $|y_2-y_1|\geqq b$, is closed. Hence the set of vertices of all the contours of the theorem, for given directions x, y, is measurable Borel. But, as we have proved, this set cannot be of zero measure spatially. This completes the theorem.

It will be noticed that Theorem I of §7.1 relates to measure and Theorem II of §7.1 is topological in character. The present theorem partakes of both characters. A set which has an exceptional point, by this theorem, cannot be of zero spatial measure, since it must contain a set of vertices of positive measure; it cannot be of dimension zero, since the neighborhood of an exceptional point must contain contours of dimension one.

In the plane, and for logarithmic potential, a boundary set s which contains an exceptional point with respect to the complement $\Sigma+B$ of s contains a family of rectangular contours, with sides parallel to given orthogonal directions, whose

vertices constitute a set of positive superficial measure. It follows then that *if s is a boundary set and occludes no points from infinity it has no exceptional points.* This is a topological theorem of which the analogue in three dimensions, as illustrated in §7.2, is not valid.

III. The energy equation

8. **The Dirichlet integral and the averaging process.** When U is a potential, the quantity

$$(\nabla U)^2 = \left(\frac{\partial U}{\partial x}\right)^2 + \left(\frac{\partial U}{\partial y}\right)^2 + \left(\frac{\partial U}{\partial z}\right)^2 \tag{1}$$

has meaning almost everywhere and is a measurable function spatially; for the separate partial derivatives, according to §3, possess these properties. Moreover the quantity in (1) is identical almost everywhere with the expression in terms of generalized or vector derivatives

$$(D_xU)^2 + (D_yU)^2 + (D_zU)^2$$

which is invariant of an orthogonal transformation at all points where D_xU, D_yU, D_zU exist. We shall discuss the convergence of the Dirichlet integrals

$$D = D(U) = \int_W (\nabla U)^2 dM, \tag{2}$$

$$D(U,V) = \int_W \left\{\frac{\partial U}{\partial x}\frac{\partial V}{\partial x} + \frac{\partial U}{\partial y}\frac{\partial V}{\partial y} + \frac{\partial U}{\partial z}\frac{\partial V}{\partial z}\right\} dM = \int_W \nabla(U,V)dM. \tag{3}$$

It is immediately verified that if $\delta(m)$ is a function such that

$$\int_E \phi^2 dM \leqq \delta(m), \qquad \int_E \psi^2 dm \leqq \delta(m)$$

when meas $E \leqq m$, then also, using the notation of §4 for the spatial average,

$$\left|\int_E \phi(\rho', M)\psi(\rho'', M)dM\right| \leqq \delta(m). \tag{4}$$

We state this fact as a lemma.

Lemma. *Let $\phi(M)$, $\psi(M)$ be summable with their squares over W. The absolute continuity of the integrals over measurable sets E,*

$$\int_E \{\phi(\rho', M)\}^2 dM, \qquad \int_E \phi(\rho', M)\psi(\rho'', M)dM,$$

is uniform as ρ' and ρ'' tend independently to zero.

If $U(M)$ is a potential of a distribution of positive mass on F, we have

$$(5) \qquad U(\rho, M) = \int_W k_\rho(M, P)df(e_P)$$

where

$$k_\rho(M, P) = \frac{3\rho^2 - MP^2}{2\rho^3}, \qquad MP < \rho,$$

$$= \frac{1}{MP}, \qquad MP \geqq \rho,$$

is continuous in M, P with continuous partial derivatives of the first order, and is superharmonic as a function of M. Moreover, since $U(\rho, M)$ vanishes continuously at ∞ and is harmonic outside a set F_ρ, of which no point is distant from F by more than ρ, it follows by §2 that $U(\rho, M)$ is itself the potential of a distribution of positive mass $f(\rho, e)$ on F_ρ; and, as is seen by integration over a sufficiently large spherical surface, the total mass is the same as that for $U(M)$. It has already been pointed out, in the proof of the theorem of §2.1 (the $U(\rho, M)$ forming an increasing sequence of potentials), that $f(\rho, e)$ converges weakly to $f(e)$ on a subsequence of values of ρ.

Since $U(M)$ has summable first partial derivatives, and satisfies (16), §3, $U(\rho, M)$ has continuous first partial derivatives and satisfies (18), §4. In particular, $D(U_\rho)$ exists. If $V(M)$ is a second potential of the same kind, $D(U_{\rho'}, V_{\rho''})$ also exists.

9. **Convergence of the Dirichlet integrals.** We shall prove the following

THEOREM. *A necessary and sufficient condition that $D(U)$ converge is that*

$$D(U_\rho) = \int_W \{\nabla U(\rho, M)\}^2 dM$$

remain bounded as ρ tends to zero. Further

$$(6) \qquad D(U) = \lim_{\rho \to 0} D(U_\rho)$$

if $D(U)$ converges.

That the condition is necessary comes immediately from the lemma of §8, applied to $\partial U_\rho/\partial x$, $\partial U_\rho/\partial y$, $\partial U_\rho/\partial z$ separately. For it follows that the absolute continuity of $\int\{\nabla U(\rho, M)\}^2 dM$ is uniform as ρ tends to zero. Hence this integral is bounded, independently of ρ, on any bounded region. But outside a certain bounded region, sufficiently large, $U(\rho, M)$ is identical with $U(M)$, for all ρ, as ρ tends to zero. Hence $\int_W\{\nabla U(\rho, M)\}^2 dM$ is bounded,

independently of ρ. Also equation (6) is verified. For since lim ($\rho=0$) $\{\nabla U(\rho, M)\}^2=\{\nabla U(M)\}^2$ for almost all M, and the absolute continuity of the integral is uniform, it follows by Vitali's theorem that

$$\lim_{\rho=0}\int_D \{\nabla U(\rho, M)\}^2 dM = \int_D \{\nabla U(M)\}^2 dM,$$

where D is any bounded region; on the other hand, the integrals extended over the portion $W-D$ of W, where D is taken sufficiently large, are identical.

The condition is also sufficient. In fact, it is well known and immediate that if $\phi_1(M)$, $\phi_2(M)$, $\cdots$ form a sequence of not negative summable functions, which have a limit, lim $(i=\infty)\phi_i(M)=\phi(M)$ almost everywhere on a perfect set E, then

$$\liminf_{i=\infty}\int_E \phi_i(M)dM \geqq \int_E \phi(M)dM, \tag{7}$$

admitting $+\infty$ as a possible value of the right hand member. Hence if K exists such that

$$\int_W \{\nabla U(\rho, M)\}^2 dM \leqq K, \text{ for all } \rho,$$

it follows from (7), taking $\rho=1/i$, that $\{\nabla U(M)\}^2$ is summable over every bounded region D, and

$$\int_D \{\nabla U(M)\}^2 dM \leqq K.$$

Hence $\{\nabla U(M)\}^2$ is summable over W.

We have also immediately the following corollary:

COROLLARY. *If $U(M)$, $V(M)$ are two potentials of positive masses on F, finite in total amount, such that $D(U)$ and $D(V)$ converge, the integral $D(U, V)$ also converges, and*

$$D(U, V) = \lim D(U_{\rho'}, V_{\rho''}), \tag{8}$$

as ρ' and ρ'' tend independently to zero.

The convergence comes immediately from the inequality $2|\nabla(U, V)| \leqq(\nabla U)^2+(\nabla V)^2$; the limit property is a consequence of the uniform absolute continuity, as in the proof of the theorem.

10. **The energy equation.** We prove the following

THEOREM. *If $D(U)$ exists, then*

$$D(U) = 4\pi \int_W U(P)df(e_P), \tag{9}$$

and if also $D(V)$ exists, and $\mu(e)$ is the mass function for V, then

$$D(U, V) = 4\pi \int_W U(P)d\mu(e_P). \tag{10}$$

Consider first (10). We remember that $U(\rho', M)$, $V(\rho'', M)$ are potentials of positive mass distributions on bounded sets, and prove the following lemma.

LEMMA. *If $D(U_{\rho'}, V_{\rho''}) = 4\pi\int_W U(\rho', P)d\mu(\rho'', e_P)$ and $D(U)$, $D(V)$ exist, then $D(U, V) = 4\pi\int_W U(P)d\mu(e_P)$.*

In fact, by the corollary of §9,

$$D(U, V) = \lim_{\rho'=0}\left\{\lim_{\rho''=0} D(U_{\rho'}, V_{\rho''})\right\} = 4\pi \lim_{\rho'=0}\left\{\lim_{\rho''=0}\int_W U(\rho', P)d\mu(\rho'', e_P)\right\}$$
$$= 4\pi \lim_{\rho'=0}\int_W U(\rho', P)d\mu(e_P)$$

by the weak convergence of $\mu(\rho'', e)$ to $\mu(e)$, over the proper sequence of values of ρ'', the function $U(\rho', M)$ being continuous. Whence, since $U(\rho', P)$ increases monotonically to $U(P)$, as ρ' tends to zero, by the definition of generalized integral,

$$D(U, V) = 4\pi \int_W U(P)d\mu(e_P),$$

which proves the lemma.

To return to the theorem, we note that $U(\rho_1', \rho_2', \rho_3', M)$, $V(\rho_1'', \rho_2'', \rho_3'', M)$ have continuous third partial derivatives, and the identity (10) applies to them as an immediate consequence of Green's theorem, their first derivatives vanishing continuously at infinity like $1/r^2$. Hence by successive applications of the lemma, the equation (10) is established for U, V. Equation (9) is obtained from (10) by writing $V = U$.

COROLLARY I. *If $D(U)$, $D(V)$ are finite, then also $D(U, V)$ is finite and positive.*

In fact, $D(U, V)$ is given by (10).

COROLLARY II. *If $D(U_1)$, $D(U_2)$, $D(V)$ are finite and $U_2 \geqq U_1$ but $U_2 \not\equiv U_1$, then $D(U_2) > D(U_1)$; $D(U_2, V) \geqq D(U_1, V)$.*

Since there is a point where $U_2 > U_1$, there will be a neighborhood of the point in which, almost everywhere, $U_2 > U_1$, by (19′), §4. But U_1, U_2 vanish continuously at ∞. Hence $D(U_2 - U_1) > 0$.

The second inequality of the theorem is an immediate consequence of (10). Moreover, since $D(U_1)$ and $D(U_2)$ are convergent Lebesgue integrals, the same is true of $D(U_2, U_1)$ and of $D(U_2 - U_1, U_1) = D(U_2, U_1) - D(U_1)$, although $U_2 - U_1$ is not necessarily a potential of positive mass; hence also $D(U_2 - U_1, U_2)$ and $D(U_2 - U_1) = D(U_2 - U_1, U_2) - D(U_2 - U_1, U_1)$ are convergent Lebesgue integrals. But

$$D(U_2) = D(U_1 + U_2 - U_1) = D(U_1) + 2D(U_2 - U_1, U_1) + D(U_2 - U_1)$$

where $D(U_2 - U_1) > 0$, and by equation (10), $D(U_2 - U_1, U_1) \geqq 0$.

Corollary III. *The integrals $D(U_{\rho'}, V_{\rho''})$, $D(U_{\rho'})$ are monotonic increasing as ρ', ρ'' decrease independently to zero.*

Corollary IV. *A necessary and sufficient condition that $D(U)$ exist is that the generalized integral $\int_W U(P)df(e_P)$ converge.*

The necessity has already been demonstrated in the theorem. For the sufficiency, we note first that

$$D(U_{\rho'}, U_{\rho''}) = 4\pi \int_W U(\rho', P)df(\rho'', e_P).$$

Choose now a sequence of values ρ_i'', decreasing to zero, such that the corresponding mass distributions $f(\rho_i'', e)$ of the $U(\rho_i'', M)$ converge in the weak sense to $f(e)$. Since $U(\rho', M)$ is continuous,

$$\lim_{\rho_i''=0} D(U_{\rho'}, U_{\rho_i''}) = 4\pi \int_W U(\rho', P)df(e_P).$$

But by Corollary III, $D(U_{\rho'}, U_{\rho''})$ increases monotonically as ρ'' decreases to zero; accordingly

$$D(U_{\rho'}, U_{\rho''}) \leqq 4\pi \int_W U(\rho', P)df(e_P) \leqq 4\pi \int_W U df, \qquad \text{for all } \rho', \rho''.$$

Hence, by taking $\rho'' = \rho$, $\rho' = \rho$, $D(U_\rho)$ is bounded independently of ρ, and $D(U)$ is finite, by the theorem of §9.

Corollary V. *A sufficient condition that $D(U)$ exist is that $U(M)$ be bounded in W.*

University of California,
Berkeley, Calif.

AN ELLIPTIC SYSTEM OF INTEGRAL EQUATIONS ON SUMMABLE FUNCTIONS*

BY

J. H. BINNEY

1. Introduction. Given $A(e)$ and $B(e)$ two additive† functions of point sets e, measurable in the Borel sense and contained in a simply connected and bounded plane open region T, we shall consider the elliptic system of integral equations

$$
(1) \qquad \begin{aligned} \int_s \phi(x, y)dy + \theta(x, y)dx &= A(\sigma), \\ \int_s - \theta(x, y)dy + \phi(x, y)dx &= B(\sigma), \end{aligned}
$$

in which s is a (variable) simple closed rectifiable curve in T and σ its interior region. In one sense it would be more natural to employ, in the right hand members of equations (1), the functions of curves which "correspond"‡ to $A(e)$ and $B(e)$, rather than the functions of point sets themselves, since the curvilinear integrals are functions of curves. But aside from the difficulty of defining additive functions of curves on a family of curves as general as rectifiable ones, it turns out that we need to consider only curves s where the functions of point sets are regular, that is, such that

$$A(\sigma) = A(\sigma + s), \qquad B(\sigma) = B(\sigma + s);$$

in fact, only curves such that if e'' is any set of points on s itself measurable Borel on s, then $A(e'') = 0$ and $B(e'') = 0$.

In this paper,§ a general solution of the pair of equations (1) is obtained, supposing merely that the functions ϕ and θ are summable superficially everywhere in T. This result is expressed in Theorem IV, below.

Let

---

* Presented to the Society, August 31, 1932; received by the editors May 31, 1934.

† "Additive" is understood as "completely additive," that is, additive over a denumerable infinity of distinct sets. An additive function is therefore bounded.

‡ Evans, *Fundamental points of potential theory*, Rice Institute Pamphlet, vol. 7 (1920), pp. 252–329. See pp. 261 and 268.

§ The author wishes to express his indebtedness for assistance in this problem to Professor G. C. Evans, who has discussed continuous solutions of the related pair of partial differential equations in *An elliptic system related to Poisson's equation*, Acta Szeged, vol. 6 (1932), pp. 27–33, and to Dr. A. J. Maria, who has also made suggestions on important points of the treatment.

$$
(2) \qquad
\begin{aligned}
\phi_0(M) &\equiv \frac{1}{2\pi}\int_T \frac{1}{MP}[\cos(MP, y)dB(e_P) - \cos(MP, x)dA(e_P)],\\
\theta_0(M) &\equiv \frac{1}{2\pi}\int_T \frac{1}{MP}[\cos(MP, x)dB(e_P) + \cos(MP, y)dA(e_P)].
\end{aligned}
$$

These functions are defined almost everywhere and are summable over the region T.* We shall see that these are particular solutions of the pair of equations (1).

2. Lemmas on convergence. It is necessary first to establish the following lemmas.

LEMMA A. *If $f(e)$ is a non-negative additive function of point sets e in T, measurable in the Borel sense, and s is a simple closed rectifiable curve in T, on which $\int_T[1/(MP)]df(e_P)$ represents a summable function of M with respect to arc length, then $f(e'')=0$ for every set e'', composed of points of s and measurable Borel on s, and the equation*

$$\int_T \frac{1}{MP}\,df(e_P) = \int_{T-s} \frac{1}{MP}\,df(e_P)$$

is valid.

The lemma is of course equally true if s is a simple rectifiable arc, and the proof which follows requires merely an obvious modification to cover this case.

At any point M of s for which $\int_T[1/(MP)]df(e_P)$ converges, we have

$$(3) \qquad \int_T \frac{1}{MP}\,df(e_P) = \int_{T-s} \frac{1}{MP}\,df(e_P') + \int_s \frac{1}{MP}\,df(e_P''),$$

where $e''=e\cdot s$ and $e'=e\cdot(T-s)$, that is, the parts of e on s and in $T-s$, respectively. These sets are measurable Borel if e is measurable Borel and $f(e)=f(e')+f(e'')$ since $e=e'+e''$ and e' and e'' have no common points. To show that (3) holds, we let

$$
\begin{aligned}
h_n(M, P) &= \frac{1}{MP}, \qquad & MP > \frac{1}{n},\\
&= n, \qquad & MP \leqq \frac{1}{n}.
\end{aligned}
$$

This is a continuous function of (M, P) and is such that $h_n(M, P)$

---

* Evans, Rice Institute Pamphlet, loc. cit. See p. 263.

$\leqq h_{n+1}(M, P)$. Since $h_n(M, P)$ is uniformly continuous as a function of P in the region $T-s$, in fact in the whole plane, we may show from the definition of the Stieltjes integral that

$$\int_T h_n(M, P)df(e_P) = \int_{T-s} h_n(M, P)df(e_P') + \int_s h_n(M, P)df(e_P'').$$

All three terms of this equation are increasing functions of n, and since the left hand member remains bounded as n becomes infinite, both terms of the right hand member do also. Hence the resulting generalized integrals converge and yield the desired equation (3).

Hence if the left hand member of equation (3) represents a summable function of M on the curve s, the same is true of the function

$$\int_s \frac{1}{MP} df(e_P'').$$

But, as is seen by direct calculation, this can be true only if $f(e'') \equiv 0$.

A well known property, which is deduced with the aid of Fubini's theorem on multiple integrals, is the following:

LEMMA B. *If $f(e)$ is an additive function of point sets e, measurable Borel in T, then the integral $\int_T [1/(MP)]df(e_P)$ exists for almost all M in T and is a summable function of M on almost all rectangular contours in T, whose sides are parallel to two fixed rectangular directions x and y.*

When the integrals in (4) below are suitably defined we may prove the following lemma.

LEMMA C. *If $f(e)$ is an additive function of point sets e in T, measurable Borel, and s is a simple closed rectifiable curve in T, on which*

$$\int_T \frac{1}{MP} | df(e_P) |$$

represents a summable function of the point M with respect to arc length, then the integrals

$$(4) \qquad \int_T \frac{1}{MP} \cos (MP, s_M)df(e_P), \qquad \int_T \frac{1}{MP} \cos (MP, n_M)df(e_P)$$

exist almost everywhere in T as summable functions of M with respect to arc length on s and the equations

$$(5) \qquad \int_s ds_M \int_T \frac{1}{MP} \cos (MP, s_M)df(e_P)$$

$$= \int_{T-s} df(e_P) \int_s \frac{1}{MP} \cos (MP, s_M) ds_M = 0,$$

$$\int_s ds_M \int_T \frac{1}{MP} \cos (MP, n_M) df(e_P)$$

(6)

$$= \int_{T-s} df(e_P) \int_s \frac{1}{MP} \cos (MP, n_M) ds_M = 2\pi f(\sigma)$$

are valid, where σ is the set of points interior to s.

In this paper we take n_M to mean the direction of the interior normal to the curve s at the point M.

We notice that $\cos (MP, s_M)$, $\cos (MP, n_M)$ are not defined when $P=M$ or when M is a point on the curve s where the direction of s is not definite. Evidently, however, we may define them in these exceptional cases so that the functions are measurable Borel in the closed space R_3 in which coordinates are $x=x_P$; $y=y_P$; $z=s_M$, s_M being measured from some fixed point on s and varying over the length of s, and $P\equiv(x,y)$ being confined to a closed rectangle R which includes the region T. We extend the definition of $f(e)$ throughout R by writing $f(e)=f(e\cdot T)$.

Since the integral over T with respect to $f(e)$ is the same as the integral over $T-s$, and since the points M on s where s has no tangent direction form a set of measure zero on s, independent of P in $T-s$, the values of the iterated integrals in (5) and (6) are independent of the definitions assigned to these cosines at the exceptional points. The lemma is then verified by a change in the order of integration which is justified by the existence of the integral*

$$\int_s ds_M \int_{T-s} \frac{1}{MP} | df(e_P) |.$$

A lemma of slightly different type is the following:

LEMMA D. *Let α, x represent any two fixed directions. Then, with the hypotheses on s and $f(e)$ of Lemma C, the integral*

$$\int_s dx \int_T \frac{1}{MP} \cos (MP, \alpha) df(e_P) = \int_s dx \int_{T-s} \frac{1}{MP} \cos (MP, \alpha) df(e_P)$$

exists, and is equal to

$$\int_{T-s} df(e_P) \int_s \frac{1}{MP} \cos (MP, \alpha) dx.$$

* G. C. Evans, Rice Institute Pamphlet, loc. cit., p. 258.

The integral in Lemma D is an iteration of two Daniell S-integrals.* For its treatment we make use of the following proposition:

LEMMA E. *If $f(t)$ is measurable Borel and if $x(t)$ is absolutely continuous in t such that $f(t)x'(t)=f(t)dx/dt$ is summable over (a, b), then $\int_a^b f(t)dx(t)$ exists and*

$$\int_a^b f(t)dx(t) = \int_a^b f(t)x'(t)dt.$$

We note that if $f(t)$ is summable and $x'(t)$ is bounded, $f(t)x'(t)$ is summable over (a, b).

We first prove the lemma for $x(t)$ monotonic-increasing. If $f(t)$ is continuous both members of the above equation exist and they are equal. With $x'(t)\geqq 0$, both members are general I-integrals of $f(t)$ (a Lebesgue integral is merely a special case of such an integral). Hence if the right hand member exists, the left member does also, and the equality is unchanged. But the right member exists if $f(t)x'(t)$ is summable.

If $x(t)$ is no longer monotonic-increasing, from the measurability of $f(t)$, $x'(t)$ and the summability of $f(t)\cdot x'(t)$ follows the summability of the functions

$$f(t)\cdot\tfrac{1}{2}\{\,|\,x'(t)\,| \pm x'(t)\}.$$

Hence if we denote by $x^+(t)$, $x^-(t)$ the positive and negative variation functions respectively for $x(t)$, with derivatives

$$x^{+\prime}(t) = \tfrac{1}{2}\{\,|\,x'(t)\,| + x'(t)\}, \qquad x^{-\prime}(t) = \tfrac{1}{2}\{\,|\,x'(t)\,| - x'(t)\},$$

the I-integrals $\int_a^b f(t)dx^+(t)=\int_a^b f(t)x^{+\prime}(t)dt$ and $\int_a^b f(t)dx^-(t)=\int_a^b f(t)x^{-\prime}(t)dt$ both exist, and $\int_a^b f(t)dx(t)$, which is an S-integral and the difference of two I-integrals, exists and has the assigned value.†

Consider now any function $Q(M)$ summable with respect to arc length s_M on the simple closed rectifiable curve s. Let us write $x_M=x(s_M)$ to represent the projection of the arc $RM=s_M$ on OX, where R is some fixed point of the curve s. Then $x(s_M)$ is an absolutely continuous function of arc length s_M. In fact,

$$x_M = x(s_M) = \int_{s_M} \cos\,(x, s_M)ds_M.$$

Hence the derivative $dx/ds_M=\cos\,(x, s_M)$ exists for almost all points M on s and is bounded in absolute value by unity. Hence by Lemma E,

* P. J. Daniell, *A general form of integral*, Annals of Mathematics, vol. 19 (1918), pp. 279–294.

† P. J. Daniell, loc. cit.

$$\int_s Q(M)dx_M \text{ exists and } = \int_s Q(M) \cos (x, s_M)ds_M.$$

We return now to the proof of Lemma D. Since the function

$$\int_T \frac{1}{MP} | df(e_P) |$$

is summable with respect to s_M on s and since it dominates the measurable function

$$\int_T \frac{1}{MP} \cos (MP, \alpha)df(e_P),$$

then the latter is summable with respect to s_M. Hence we may take

$$Q(M) = \int_T \frac{1}{MP} \cos (MP, \alpha)df(e_P).$$

Therefore

$$\int_s dx_M \int_T \frac{1}{MP} \cos (MP, \alpha)df(e_P)$$
$$= \int_s ds_M \int_T \frac{1}{MP} \cos (MP, \alpha) \cos (x, s_M)df(e_P).$$

But now the change of order of integration may be justified in the same manner as in Lemma C. Also $f(e \cdot s) = 0$ and our integral is therefore

$$\int_{T-s} df(e_P) \int_s \frac{1}{MP} \cos (MP, \alpha) \cos (x, s_M)ds_M$$

and this again is

$$\int_{T-s} df(e_P) \int_s \frac{1}{MP} \cos (MP, \alpha)dx_M.$$

This completes the proof of Lemma D.

3. The general solution of equations (1). We are now in a position to prove the following theorem.

THEOREM I. *The functions ϕ_0 and θ_0 given by* (2) *satisfy the pair of equations* (1) *on all simple closed rectifiable curves s in T on which*

$$\int_T \frac{1}{MP} | dA(e_P) |, \qquad \int_T \frac{1}{MP} | dB(e_P) |$$

represent summable functions of M with respect to arc length on s.

Since the functions

$$\int_T \frac{1}{MP} \mid dA(e_P) \mid, \qquad \int_T \frac{1}{MP} \mid dB(e_P) \mid$$

are summable with respect to arc length on s and since x and y are absolutely continuous functions of s_M with bounded derivatives, we can conclude from Lemma E, as in Lemma D, that the integral

$$\int_s \phi_0 dy + \theta_0 dx$$

exists and, moreover, that

$$\int_s \phi_0 dy + \theta_0 dx = \int_s \left\{ \phi_0 \frac{dy}{ds_M} + \theta_0 \frac{dx}{ds_M} \right\} ds_M.$$

Hence we are justified in substituting ϕ_0 and θ_0 in the left hand member of the first of the equations (1). Doing this and noting that, almost everywhere on s,

$$\frac{dy}{ds_M} = \cos(y, s_M) = -\cos(x, n_M),$$

$$\frac{dx}{ds_M} = \cos(x, s_M) = \cos(y, n_M),$$

we get

$$\int_s \phi_0 dy + \theta_0 dx = \int_s \{ -\phi_0 \cos(x, n_M) + \theta_0 \cos(y, n_M) \} ds_M.$$

But the right hand member may be written as follows, since $A(e \cdot s) \equiv 0 \equiv B(e \cdot s)$:

$$\begin{aligned} \frac{1}{2\pi} \int_s ds_M \int_{T-s} \frac{1}{MP} \{ & -\cos(MP, y)\cos(x, n_M) dB(e_P) \\ & + \cos(MP, x)\cos(x, n_M) dA(e_P) + \cos(MP, x)\cos(y, n_M) dB(e_P) \\ & + \cos(MP, y)\cos(y, n_M) dA(e_P) \} \\ = \frac{1}{2\pi} \int_s ds_M \int_{T-s} \frac{1}{MP} \{ & [\cos(MP, x)\cos(y, n_M) \\ & - \cos(MP, y)\cos(x, n_M)] dB(e_P) + [\cos(MP, x)\cos(x, n_M) \\ & + \cos(MP, y)\cos(y, n_M)] dA(e_P) \} \end{aligned}$$

$$= \frac{1}{2\pi}\int_s ds_M \int_{T-s} \frac{1}{MP}\{[\cos(MP,x)\cos(x,s_M)$$
$$+ \cos(MP,y)\cos(y,s_M)]dB(e_P) + [\cos(MP,x)\cos(x,n_M)$$
$$+ \cos(MP,y)\cos(y,n_M)]dA(e_P)\}.$$

We have, for almost all M on s,

$$\cos(MP,x)\cos(x,n_M) + \cos(MP,y)\cos(y,n_M) = \cos(MP,n_M),$$
$$\cos(MP,x)\cos(x,s_M) + \cos(MP,y)\cos(y,s_M) = \cos(MP,s_M).$$

Hence for almost all M on s, the inside integrals take the desired form

$$\int_{T-s}\frac{1}{MP}\cos(MP,s_M)dB(e_P), \qquad \int_{T-s}\frac{1}{MP}\cos(MP,n_M)dA(e_P).$$

Hence the Lebesgue integral with respect to s_M of the inside integrals has the value

$$\frac{1}{2\pi}\int_s ds_M\int_{T-s}\frac{1}{MP}\cos(MP,s_M)dB(e_P)$$
$$+\frac{1}{2\pi}\int_s ds_M\int_{T-s}\frac{1}{MP}\cos(MP,n_M)dA(e_P).$$

By making use now of the results of Lemma C we see that

$$\int_s \phi_0 dy + \theta_0 dx = A(\sigma)$$

and ϕ_0, θ_0 satisfy the first equation of (1). In a similar manner we can show that ϕ_0 and θ_0 also satisfy the second equation of (1) for s.

From Lemma B and Theorem I, we have the following corollary.

COROLLARY. *The functions ϕ_0 and θ_0 given by* (2) *satisfy the pair of equations* (1) *on almost all rectangles in T, with sides parallel to x and y.*

These results lead immediately to the following theorem:

THEOREM II. *If the functions $\bar{\phi}$ and $\bar{\theta}$ are solutions of equations* (1) *on almost all rectangles in T, the functions*

$$\phi = \bar{\phi} - \phi_0, \qquad \theta = \bar{\theta} - \theta_0$$

satisfy the pair of equations

$$(7) \qquad \int_s \phi dy + \theta dx = 0, \qquad \int_s -\theta dy + \phi dx = 0,$$

on almost all rectangles in T, ϕ_0 and θ_0 being given by (2).

We state the following theorem.*

THEOREM III. *If ϕ and θ are two functions summable superficially over every closed region interior to T and satisfy equations* (7) *on almost all rectangles in T, then there exists a function $\Psi(M)$ harmonic in T, such that*

$$\frac{\partial\Psi}{\partial x} = \theta(M), \qquad \frac{\partial\Psi}{\partial y} = \phi(M)$$

for almost all M in T.

Finally, we may summarize the results of the preceding pages in the following theorem:

THEOREM IV. *The functions*

$$\phi \equiv \phi_0 + \frac{\partial\psi}{\partial y}; \qquad \theta \equiv \theta_0 + \frac{\partial\psi}{\partial x} \tag{8}$$

where ϕ_0 and θ_0 are given by (2) *and where ψ is an arbitrary solution of Laplace's equation in T, form a system of solutions of the pair of equations* (1) *for all simple closed rectifiable curves s, interior to the region T, on which the integrals*

$$\int_T \frac{1}{MP}\,|\,dA(e_P)\,|, \qquad \int_T \frac{1}{MP}\,|\,dB(e_P)\,|$$

represent summable functions of M on s, with respect to ds_M; in particular the functions (8) *are solutions of* (1) *for almost all rectangular contours interior to T with sides parallel to given directions x and y.*

Conversely, if ϕ and θ are summable superficially over any closed region interior to T and are solutions of the pair of equations (1) *on almost all rectangles, interior to the region T, with sides parallel to the axes, they may be expressed in the form* (8) *in $T-E(x, y)$, where $E(x, y)$ is a set of points of T having superficial measure zero at most.*

4. The complex plane. It is interesting to note that we might have begun our study of the pair of equations (1) by considering a single equation in complex variables.

---

* The editors of these Transactions have brought to the author's attention the fact that substantially this theorem is proved by V. S. Fedoroff, *Sur le théorème de Morera*, Moscow Mathematicheskii Sbornik, vol. 40 (1933), pp. 168–179. Consequently it is unnecessary to give a proof of Theorem III. See also G. C. Evans, *Note on a theorem of Bôcher*, American Journal of Mathematics, vol. 50 (1928), pp. 123–126. For Theorem III a function Ψ is constructed which is a "potential of its generalized derivatives" θ, ϕ, and satisfies the second of (7); Ψ is then harmonic by the generalized theorem of Bôcher.

In order to show the connection of this pair of equations with a single equation in the complex plane, we let $f(z)$ be a function of the complex variable $z=x+iy$ such that its real and its imaginary parts are summable superficially over any closed region which is interior to a simply connected plane bounded open region T. Let $\Phi(e)$ be a completely additive complex-valued function of point sets e in T measurable in the Borel sense. We now consider the equation

$$\int_s f(z)dz = \Phi(\sigma) \tag{9}$$

where s is a simple closed rectifiable curve in the region T and σ is the set of interior points of s. If we place

$$\Phi(e) = B(e) + iA(e)$$

and let

$$f(z) = \phi(x, y) + i\theta(x, y),$$

equation (9) becomes

$$\int_s (\phi + i\theta)(dx + idy) = B(\sigma) + iA(\sigma).$$

If we now equate the real and the imaginary parts we obtain the pair of equations

$$\begin{aligned} \int_s \phi dy + \theta dx &= A(\sigma), \\ \int_s - \theta dy + \phi dx &= B(\sigma) \end{aligned} \tag{1}$$

which we have been studying. We may accordingly state the following theorem:

THEOREM V. *If s is a simple closed rectifiable curve in T on which the integrals*

$$\int_T \frac{1}{MP}\,|\,dA(e_P)\,|, \qquad \int_T \frac{1}{MP}\,|\,dB(e_P)\,|$$

represent summable functions of M with respect to arc length s_M, then

$$f(z) \equiv F(z) + \frac{1}{2\pi}\int_T \frac{1}{MP}\, e^{i\alpha} d\Phi(e_P) \tag{10}$$

is a solution of equation (9), *where* α *is the angle* (MP, y) *and* $F(z)$ *is an arbitrary holomorphic function in* T.

Since

$$\int_s F(z)dz = 0$$

by the Cauchy integral theorem, it only remains to show that

$$\frac{1}{2\pi}\int_T \frac{1}{MP}\, e^{i\alpha}d\Phi(e_P)$$

satisfies equation (9). And this is verified by direct calculation.

As a result of this theorem and of Lemma B we have the following corollary.

COROLLARY. *The function*

$$f(z) \equiv F(z) + \frac{1}{2\pi}\int_T \frac{1}{MP}\, e^{i\alpha}d\Phi(e_P)$$

is a solution of equation (9) *on almost all rectangles in* T.

We also have a converse result expressed by the following theorem.

THEOREM VI. *If* ϕ *and* θ *are two functions summable over every closed region interior to* T *and if* $\phi+i\theta$ *is a solution of equation* (9) *on almost all rectangular contours* s *in* T *whose sides are parallel to the given directions* x *and* y, *then* $\phi+i\theta$ *may be expressed in the form* (10) *in* $T-E(x, y)$ *where* $E(x, y)$ *is a set of points of* T *of superficial measure zero.*

Since $\phi+i\theta$ is a solution of the equation (9) on almost all rectangles s in T we have

$$\int_s (\phi + i\theta)dz = \Phi(\sigma). \tag{11}$$

In fact the functions

$$u = \phi - \phi_0, \qquad v = \theta - \theta_0$$

satisfy the equations

$$\int_s udy + vdx = 0, \qquad \int_s - vdy + udx = 0$$

on almost all rectangles in T. Hence, almost everywhere in T,

$$u = \frac{\partial \Psi}{\partial y}, \qquad v = \frac{\partial \Psi}{\partial x},$$

where Ψ is harmonic in T. But then the function

$$F(z) = u + iv = \frac{\partial \Psi}{\partial y} + i\,\frac{\partial \Psi}{\partial x}$$

is holomorphic in T.

RICE INSTITUTE,
HOUSTON, TEXAS

ON CONVEX FUNCTIONS*

BY

TIBOR RADÓ

Introduction

0.1. The problems raised and solved in this paper were suggested by certain results concerning subharmonic functions, published jointly by Dr. E. F. Beckenbach and myself.†

0.2. According to F. Riesz, a function $u(x, y)$, continuous in a certain region R of the xy-plane, is called *subharmonic* if the inequality

$$u(x, y) \leqq \frac{1}{2\pi}\int_0^{2\pi} u(x + \rho\cos\theta, y + \rho\sin\theta)d\theta$$

holds for every point (x, y) in R and for every value of $\rho > 0$, such that the circular disc with center (x, y) and radius ρ is entirely comprised in R.‡

0.3. We have the following theorem:

A function $u(x, y)$, continuous in a region R, is subharmonic there if and only if the inequality

$$\frac{1}{\rho^2\pi}\iint_{\xi^2+\eta^2<\rho^2} u(x + \xi, y + \eta)d\xi d\eta \leqq \frac{1}{2\pi}\int_0^{2\pi} u(x + \rho\cos\theta, y + \rho\sin\theta)d\theta$$

holds for every point (x, y) in R and for every $\rho > 0$, such that the circular disc with center (x, y) and radius ρ is entirely comprised in R.§

0.4. In the second of the joint papers by Dr. Beckenbach and myself, referred to in 0.1, the following theorem was proved in connection with an investigation of the isoperimetric inequality.

The logarithm of a function $u(x, y)$, continuous and positive in a region R, is subharmonic there if and only if the inequality

$$\left[\frac{1}{\rho^2\pi}\iint_{\xi^2+\eta^2<\rho^2} u(x + \xi, y + \eta)^2 d\xi d\eta\right]^{1/2} \leqq \frac{1}{2\pi}\int_0^{2\pi} u(x + \rho\cos\theta, y + \rho\sin\theta)d\theta$$

* Presented to the Society, April 7, 1934; received by the editors June 7, 1934.

† *Subharmonic functions and minimal surfaces*, these Transactions, vol. 35 (1933), pp. 648–661. *Subharmonic functions and surfaces of negative curvature*, these Transactions, vol. 35 (1933), pp. 662–674.

‡ See F. Riesz, *Sur les fonctions subharmoniques* etc., in two parts, Acta Mathematica, vol. 48 1926), pp. 329–343, and vol. 54 (1930), pp. 321–360.

§ See the second paper quoted under † above.

holds for every point (x, y) *in* R *and for every* $\rho > 0$, *such that the circular disc with center at* (x, y) *and with radius* ρ *is entirely comprised in* R.

0.5. The theorems in 0.3 and 0.4 are clearly two links in a chain of similar theorems. In the paper referred to in 0.4, some rather incomplete remarks were made concerning the character of these theorems. On account of the close analogy between subharmonic functions of two variables on the one hand and convex functions of one variable on the other,† there arose in this manner certain problems concerned with convex functions. We shall indicate briefly the character of the problems thus suggested.

0.6. Let $f(x)$ be a function, continuous and positive in a given open interval $x_1 < x < x_2$. Denoting by α a real exponent, we define

$$I(f, x, h, \alpha) = \left[\frac{1}{2h}\int_{-h}^{h} f(x+\xi)^{\alpha} d\xi\right]^{1/\alpha}, \text{ if } \alpha \neq 0,$$

and

$$I(f, x, h, 0) = \exp\left[\frac{1}{2h}\int_{-h}^{h} \log f(x+\xi) d\xi\right],$$

where x and h are supposed to satisfy the inequalities

$$x_1 < x - h < x + h < x_2.$$

Then $I(f, x, h, \alpha)$ is a continuous and increasing function of α, for $-\infty < \alpha < +\infty$.‡

0.7. Given a real exponent γ, we define the class C_γ as the class of all functions $f(x)$ which are continuous and positive in $x_1 < x < x_2$ and which are such that

$$f^{\gamma} \operatorname{sgn} \gamma \text{ is convex, if } \gamma \neq 0,$$

and

$$\log f \text{ is convex, if } \gamma = 0.$$

0.8. Given a real exponent δ, we define the class C_δ^* as the class of all functions $f(x)$ which are continuous and positive in $x_1 < x < x_2$ and which satisfy the inequality

$$I(f, x, h, \delta) \leqq \frac{f(x-h) + f(x+h)}{2}$$

for all values of x and h such that $x_1 < x - h < x + h < x_2$.

† See, for instance, P. Montel, *Sur les fonctions convexes et les fonctions sousharmoniques*, Journal de Mathématiques, (9), vol. 7 (1928), pp. 29–60.

‡ See, for instance, Pólya and Szegö, *Aufgaben und Lehrsätze aus der Analysis*, vol. 1 (Berlin, Springer, 1925), problem 82 on p. 54 and problem 83 on p. 55.

0.9. The theorems in 0.3 and 0.4 suggest, by implication, the problem of determining couples of exponents (γ, δ) such that $C_\gamma \equiv C_\delta^*$. In conversations between Dr. Beckenbach and myself, it was found† that

$$C_\delta^* \subset C_\gamma \text{ for } 2\gamma + \delta - 3 = 0.‡$$

Dr. Beckenbach proved then that

$$C_\gamma \subset C_\delta^* \text{ for } 2\gamma + \delta - 3 = 0 \text{ and } 0 \leqq \gamma \leqq 1.$$

0.10. Thus it was established that

$$C_\gamma \equiv C_\delta^* \text{ for } 2\gamma + \delta - 3 = 0 \text{ and } 0 \leqq \gamma \leqq 1.‡$$

In a seminar on convex functions, held at Ohio State University in 1933–34, the topics just described came up again for discussion. I found, against my own expectation, that the implication

$$C_\gamma \subset C_\delta^* \text{ for } 2\gamma + \delta - 3 = 0 \tag{1}$$

did not hold generally. On the other hand, members of that seminar verified (1) for a number of values of γ outside of the interval $0 \leqq \gamma \leqq 1$, considered by Dr. Beckenbach. The purpose of this paper is to present the results of an investigation suggested by the situation just described.

0.11. For the sake of brevity, we restrict ourselves to state in this introduction the main results only. A number of applications, in particular a complete discussion of the relation $C_\gamma \equiv C_\delta^*$, will be considered in §4.

0.12. Besides the mean $I(f, x, h, \alpha)$, defined in 0.6, we shall use also another mean value $A(f, x, h, \beta)$, defined as follows:

$$A(f, x, h, \beta) = \left[\frac{f(x-h)^\beta + f(x+h)^\beta}{2}\right]^{1/\beta}, \text{ if } \beta \neq 0,$$

and

$$A(f, x, h, 0) = \exp\left[\frac{\log f(x-h) + \log f(x+h)}{2}\right] = [f(x-h)f(x+h)]^{1/2}.$$

Again, β is a real exponent, $f(x)$ a function continuous and positive in $x_1 < x < x_2$, and x and h are supposed to satisfy $x_1 < x - h < x + h < x_2$. As is

† These conversations, as well as the investigations published in the papers quoted in foot note †, p. 266, were carried on in 1932–33 while Dr. Beckenbach worked as a National Research Fellow at Ohio State University.

‡ For the sake of accuracy, it should be mentioned that we only considered at that time the case of functions with continuous first and second derivatives. Concerning the case of general continuous functions, see the remarks made in 0.15.

well known, $A(f, x, h, \beta)$ is a continuous and increasing function of β for $-\infty<\beta<+\infty$.†

0.13. *Denote by E the set of all pairs* (α, β) *for which the following assertion is true: every function* $f(x)$, *which is continuous, positive and convex in* $x_1<x<x_2$ *satisfies the inequality* $I(f, x, h, \alpha)\leqq A(f, x, h, \beta)$ *for all values of x and h such that* $x_1<x-h<x+h<x_2$.

Denote by $\overline{E}$ *the set of all pairs* (α, β) *for which the following assertion is true: if a function* $f(x)$, *which is positive and continuous in* $x_1<x<x_2$, *satisfies the inequality* $I(f, x, h, \alpha)\leqq A(f, x, h, \beta)$ *for all values of x and h such that* $x_1<x-h<x+h<x_2$, *then* $f(x)$ *is convex in* $x_1<x<x_2$.

Our main result is the explicit determination of these sets E, $\overline{E}$. As we shall prove in §2, a pair (α, β) *belongs to E if and only if one of the following four conditions is satisfied.*‡

I. $\alpha \leqq -2$ *and* $\beta \geqq 0$.

II. $-2\leqq\alpha\leqq-\dfrac{1}{2}$ *and* $\beta\geqq\dfrac{\alpha+2}{3}\cdot$

III. $-\dfrac{1}{2}\leqq\alpha\leqq 1$ *and* $\beta\geqq\dfrac{\alpha\log 2}{\log(\alpha+1)}\cdot$

IV. $1\leqq\alpha$ *and* $\beta\geqq\dfrac{\alpha+2}{3}\cdot$

As we shall prove in §3, a pair (α, β) *belongs to* $\overline{E}$ *if and only if* $3\beta-\alpha-2\leqq 0$.

The corresponding results for concave functions will be given in §4.

0.14. The following remarks should be made concerning the method used in this paper. Whenever we shall be concerned with deriving an inequality of the form $I(f, x, h, \alpha)\leqq A(f, x, h, \beta)$ for a function $f(x)$ with certain convexity

---

† See second footnote on p. 267.

‡ The following remarks may be helpful to the reader in making a picture of the set E. Define three functions $\psi_1(\alpha)$, $\psi_2(\alpha)$, $\psi_3(\alpha)$ as follows:

$$\psi_1(\alpha)\equiv 0,\quad -\infty<\alpha<+\infty;$$

$$\psi_2(\alpha)=\frac{\alpha+2}{3},\quad -\infty<\alpha<+\infty;$$

$$\psi_3(\alpha)=\begin{cases}0, \text{ if } -\infty<\alpha\leqq-1,\\ \dfrac{\alpha\log 2}{\log(\alpha+1)}, \text{ if } -1<\alpha<+\infty \text{ and } \alpha\neq 0,\\ \log 2, \text{ if } \alpha=0.\end{cases}$$

Then E consists of all points (α, β) such that

$$\beta\geqq\max\left[\psi_1(\alpha), \psi_2(\alpha), \psi_3(\alpha)\right],$$

where the symbol max means the largest of the numbers which it precedes. The function $\psi_3(\alpha)$ is positive, increasing and concave for $-1<\alpha<+\infty$, as is easily seen by differentiation. The curves $\beta=\psi_2(\alpha)$ and $\beta=\psi_3(\alpha)$ intersect each other at the points $(-2, 0)$, $(-\frac{1}{2}, \frac{1}{2})$ and $(1, 1)$.

properties, it will be possible to reduce the discussion to the case when $f(x)$ is linear. The means I and A can be computed then explicitly, and our arguments will consist, generally speaking, of a more or less accurate discussion of the graphs of certain elementary functions. The result

$$(2) \qquad C_\gamma \subset C_\delta^* \text{ for } 2\gamma + \delta - 3 = 0 \text{ and } 0 \leqq \gamma \leqq 1,$$

proved by Dr. Beckenbach, is a particular case which comes under this description (see 4.8–4.9 below). The proof of Dr. Beckenbach for (2) consisted of quite elementary, but rather elaborate, computations, and the same remark applies, unfortunately to an even greater extent, to the arguments used in the present paper. While the general appearance of our inequalities strongly suggests the use of the simple and fundamental inequalities named after Hölder and Minkowski,† I was unable to establish a connection in this direction.

0.15. Whenever we shall be concerned with deriving, from an inequality of the form $I(f, x, h, \alpha) \leqq A(f, x, h, \beta)$, certain properties of convexity for $f(x)$, the results will be quite trivial in case $f(x)$ has continuous first and second derivatives. On the other hand, I had to follow rather devious ways to deal with the case of general continuous functions.‡ While it seems quite natural to establish those results first under the assumption of continuous first and second derivatives and to handle the general case by approximation, I was unable to carry out this program in a generality sufficient for the purposes of this paper.

The purpose of the remarks made in 0.14 and 0.15 is to call the attention of the reader to situations which might quite possibly suggest some interesting investigations.

0.16. In Part 1 of this paper, certain elementary functions, used in the sequel, are discussed. In Parts 2 and 3 respectively, we shall give the explicit determination of the sets E and $\overline{E}$, defined in 0.13. Finally, Part 4 will be concerned with extensions and miscellaneous applications.

1. Lemmas

1.1. Let $\omega_1, \omega_2, \cdots, \omega_n$ and $a_1, a_2, \cdots, a_n$ be real numbers, and consider the function

$$Q(u) = a_1 u^{\omega_1} + a_2 u^{\omega_2} + \cdots + a_n u^{\omega_n}$$

of the variable $u > 0$. It is assumed that not all the a's vanish. Then the num-

† See the very elegant presentation by F. Riesz, *Su alcune disuguaglianze*, Bollettino della Unione Matematica Italiana, vol. 7 (1928), pp. 77–79.

‡ See 3.4 to 3.10.

ber of positive roots of the equation $Q(u)=0$ is $\leqq n-1$, every root being counted with the proper multiplicity.†

1.2. Let α, β denote real numbers such that

$$\alpha \neq 0,\ \alpha+1 \neq 0,\ \beta \neq 0,\ 3\beta-\alpha-2 \neq 0, \tag{3}$$

and consider the function

$$P(u) = -\frac{1}{\alpha}u^{\alpha+\beta+1} + u^{\alpha+2} - \frac{\alpha+1}{\alpha}u^{\alpha+1} + \frac{\alpha+1}{\alpha}u^{\beta+1} - u^{\beta} + \frac{1}{\alpha}u,$$

of the real variable $u>0$. We shall need a few simple properties of $P(u)$.

1.3. Given α and β, such that (3) is satisfied, there exists an $\epsilon>0$ such that

$$\operatorname{sgn} P(u) = \operatorname{sgn}\left[(\alpha+1)(3\beta-\alpha-2)\right] \text{ for } 1-\epsilon < u < 1.$$

This follows immediately from

$$P(1) = P'(1) = P''(1) = 0,\ P'''(1) = -(\alpha+1)(3\beta-\alpha-2),$$

by using Taylor's formula.

1.4. The function $P(u)$, defined in 1.2, has at most one root in the interval $0<u<1$. If such a root exists, then it is a simple root.

Indeed, on account of $P(1)=P'(1)=P''(1)=0$, $P'''(1)\neq 0$, we have a root of multiplicity 3 at $u=1$. A direct computation shows that

$$u^{\alpha+\beta+2}P\left(\frac{1}{u}\right) = -P(u).$$

Hence, if there are σ roots in the interval $0<u<1$, then there are also σ roots in the interval $1<u<+\infty$. By 1.1, there are at most five positive roots. Hence, $3+2\sigma\leqq 5$, and consequently $\sigma\leqq 1$.

1.5. Supposing again that α, β satisfy conditions (3), we shall consider the function

$$\phi(u) = \log\frac{\left[\dfrac{1}{\alpha+1}\,\dfrac{u^{\alpha+1}-1}{u-1}\right]^{1/\alpha}}{\left(\dfrac{1+u^{\beta}}{2}\right)^{1/\beta}},\quad 0<u<1. \tag{4}$$

We shall need a few simple properties of this function $\phi(u)$.

1.6. Given α and β, such that (3) is satisfied, there exists an $\epsilon>0$ such that

$$\operatorname{sgn}\phi(u) = -\operatorname{sgn}(3\beta-\alpha-2) \text{ for } 1-\epsilon < u < 1. \tag{5}$$

Indeed, we see from (4) that

† The change of variable $u=e^{x}$ reduces $Q(u)$ to an exponential polynomial for which the result is well known. Cf. J. Tamarkin, *Some general problems*, etc., Mathematische Zeitschrift, vol. 27 (1928), p. 28.

(6) $$\phi(u) \to 0 \text{ for } u \to 1.$$

We also find, by a direct computation, that

(7) $$(1 + u^{\beta})(u^{\alpha+1} - 1)(u - 1)u\phi'(u) = P(u),$$

where $P(u)$ is the function defined in 1.2. Since $0<u<1$, it follows that

(8) $$\operatorname{sgn} \phi'(u) = \operatorname{sgn} [(\alpha + 1)P(u)], \; 0 < u < 1.$$

By 1.3 there exists an $\epsilon>0$ such that

(9) $$\operatorname{sgn} \phi'(u) = \operatorname{sgn} (3\beta - \alpha - 2) \text{ for } 1 - \epsilon < u < 1.$$

Relation (5) is derived from (6) and (9) by using the mean-value theorem.

1.7. By inspection of formula (4) it is seen that

$$\phi(+0) < 0 \text{ if } \begin{cases} \alpha + 1 < 0 \text{ and } \beta > 0, \text{ or} \\ \alpha + 1 > 0 \text{ and } \beta > \dfrac{\alpha \log 2}{\log (\alpha + 1)} \cdot \end{cases}$$

On the other hand,

$$\phi(+0) > 0 \text{ if } \begin{cases} \alpha + 1 < 0 \text{ and } \beta < 0, \text{ or} \\ \alpha + 1 > 0 \text{ and } \beta < \dfrac{\alpha \log 2}{\log (\alpha + 1)} \cdot \end{cases}$$

1.8. If $3\beta-\alpha-2>0$, then

(10) $$\phi(u) < \max [\phi(+0), 0] \text{ for } 0 < u < 1.$$

Indeed, we have (see formula (9) in 1.6),

$$\operatorname{sgn} \phi'(u) = \operatorname{sgn} (3\beta - \alpha - 2) \text{ for } 1 - \epsilon < u < 1,$$

where ϵ is some positive constant. Since, by assumption, $3\beta-\alpha-2>0$, we have therefore

(11) $$\phi'(u) > 0 \text{ for } 1 - \epsilon < u < 1.$$

Case I. $\phi'(u) \neq 0$ in $0<u<1$. By (11) $\phi'(u)>0$ in $0<u<1$. Hence $\phi(u)$ is increasing in $0<u<1$. Also, $\phi(u)\to 0$ for $u\to 1$. Thus $\phi(u)<0$ in $0<u<1$, and (10) is proved.

Case II. $\phi'(u)$ has some zero in $0<u<1$. From (7) it follows that $\phi'(u)$ and $P(u)$ have the same number of roots in $0<u<1$. But (see 1.4) $P(u)$ has at most one root in $0<u<1$. Thus $\phi'(u)$ has exactly one root in $0<u<1$, and this root, which we denote by u_0, is a simple root. Hence $\phi'(u)$ changes its sign at u_0. Since $\phi'(u)>0$ for u close to 1 (see (11)),

$$\phi'(u) < 0 \text{ for } 0 < u < u_0,$$

and

$$\phi'(u) > 0 \text{ for } u_0 < u < 1.$$

That is to say, $\phi(u)$ increases in $u_0<u<1$ and decreases in $0<u<u_0$. Consequently,

$$\phi(u) < \phi(+0) \text{ in } 0 < u \leqq u_0,$$

and

$$\phi(u) < \lim_{u\to 1} \phi(u) = 0 \text{ in } u_0 \leqq u < 1.$$

Thus (10) is proved.

1.9. In a similar way, we obtain the following lemma:

If $3\beta-\alpha-2<0$, *then*

$$\phi(u) > \min\,[\phi(+0), 0] \textit{ for } 0 < u < 1.$$

2. Determination of the set E

2.1. In 0.13, we defined the set E as the set of all pairs (α, β) such that the inequality

$$I(f, x, h, \alpha) \leqq A(f, x, h, \beta) \tag{12}$$

is satisfied by every function $f(x)$ which is positive, continuous and convex in $x_1<x<x_2$. The inequality (12) is supposed to be satisfied for all values of x and h such that $x_1<x-h<x+h<x_2$. We shall determine this set E explicitly.

2.2. Suppose that (α, β) satisfies one of the conditions I–IV of 0.13. Then, as we shall show presently, (α, β) is in E.

2.3. We first prove this under the assumption that $f(x)$ is linear:†

$$f(x) = l(x) = ax + b > 0 \text{ in } x_1 < x < x_2,$$

and

$$\alpha \neq 0,\ \alpha + 1 \neq 0,\ \beta \neq 0,\ 3\beta - \alpha - 2 \neq 0,\ \beta \neq \frac{\alpha \log 2}{\log(\alpha+1)}. \tag{13}$$

Without loss of generality, we can assume that $l(x-h)\neq l(x+h)$. Otherwise, $l(x)$ reduces to a constant and (12) is trivial.

We write the inequality (12) in the form

$$\log \frac{I(f, x, h, \alpha)}{A(f, x, h, \beta)} \leqq 0. \tag{14}$$

Since $f(x)$ has the special form $ax+b$, the mean $I(f, x, h, \alpha)$ can be computed explicitly. Thus we find that (14) is equivalent to the inequality

† Cf. 0.14.

$$\phi(u) \leqq 0 \text{ for } 0 < u < 1,$$

where $\phi(u)$ is the function defined in 1.5 and

$$u = \frac{\min\,[l(x-h),\, l(x+h)]}{\max\,[l(x-h),\, l(x+h)]}.$$

Using 1.7, we see that in every one of the four cases stated in 2.2 we have, with regard to (13),

$$\phi(+0) < 0$$

and

$$3\beta - \alpha - 2 > 0.$$

Hence (see 1.8)

$$\phi(u) < \max\,[\phi(+0), 0] = 0 \text{ in } 0 < u < 1.$$

Thus our assertion is proved under the assumption that $f(x)$ is linear and that α, β are further restricted by (13).

2.4. Let now $f(x)$ be a general continuous and positive convex function in $x_1 < x < x_2$. Let (α, β) satisfy one of the four conditions of 0.13, and suppose for a moment that (13) is also satisfied. Let x and h be such that $x_1 < x-h < x+h < x_2$, and denote by l the linear function which coincides with f at $x-h$ and $x+h$. Then we have

$$A(f, x, h, \beta) = A(l, x, h, \beta). \tag{15}$$

Since f is convex, we have $f \leqq l$ in the interval $(x-h, x+h)$. Hence:

$$I(f, x, h, \alpha) \leqq I(l, x, h, \alpha). \tag{16}$$

By 2.3 we have

$$I(l, x, h, \alpha) \leqq A(l, x, h, \beta). \tag{17}$$

(15), (16) and (17) yield

$$I(f, x, h, \alpha) \leqq A(f, x, h, \beta). \tag{18}$$

Thus the assertion in 2.2 is proved under the restriction (13). An easy continuity consideration will allow us, however, to remove this restriction and thus to complete the proof of our assertion in 2.2.

2.5. So far we have proved that every pair (α, β) which satisfies one of the four conditions of 0.13 belongs to E. We shall prove now that every pair (α, β) in E satisfies one of the four conditions 0.13 and then the theorem of 0.13 concerning E will be completely proved.

2.6. We shall need the following trivial remark: if $(\alpha, \beta) \subset E$, and $\eta \geqq 0$, then $(\alpha - \eta, \beta + \eta) \subset E$.

To see this, let $f(x)$ be any positive and continuous convex function in $x_1<x<x_2$, and let x and h be such that $x_1<x-h<x+h<x_2$. Since $(\alpha, \beta) \subset E$, we have then

$$I(f, x, h, \alpha) \leqq A(f, x, h, \beta). \tag{19}$$

We also have, since I and A are increasing functions of their last arguments, the inequalities

$$I(f, x, h, \alpha - \eta) \leqq I(f, x, h, \alpha), \tag{20}$$

$$A(f, x, h, \beta) \leqq A(f, x, h, \beta + \eta). \tag{21}$$

(19), (20) and (21) yield

$$I(f, x, h, \alpha - \eta) \leqq A(f, x, h, \beta + \eta).$$

Thus $(\alpha-\eta, \beta+\eta) \subset E$.

2.7. Suppose now that $(\alpha, \beta) \subset E$, and suppose that, in contradiction to the assertion made in 2.5, the pair (α, β) satisfies none of the four conditions of 0.13. We shall show that these assumptions lead to a contradiction.

Let η be a small positive number, and put

$$\bar{\alpha} = \alpha - \eta, \ \bar{\beta} = \beta + \eta.$$

Then, if $\eta>0$ is sufficiently small, $(\bar{\alpha}, \bar{\beta})$ will not satisfy any of the four conditions of 0.13 either. Furthermore, if $\eta>0$ is suitably chosen, $(\bar{\alpha}, \bar{\beta})$ will satisfy the relations

$$\bar{\alpha} \neq 0, \ \bar{\alpha} + 1 \neq 0, \ \bar{\beta} \neq 0, \ 3\bar{\beta} - \bar{\alpha} - 2 \neq 0.$$

Finally, on account of 2.6, $(\bar{\alpha}, \bar{\beta})$ will also belong to E.

2.8. Writing again (α, β) for $(\bar{\alpha}, \bar{\beta})$ to simplify the notation, we would have a pair (α, β) satisfying the following conditions:

$$\begin{gathered} (\alpha, \beta) \subset E; \\ \alpha \neq 0, \ \alpha + 1 \neq 0, \ \beta \neq 0, \ 3\beta - \alpha - 2 \neq 0. \end{gathered} \tag{22}$$

Besides, (α, β) satisfies none of the four conditions of 2.2. Clearly, (α, β) satisfies then one of the following three conditions:

$$\alpha + 1 < 0 \text{ and } \beta < 0, \tag{23}$$

$$\alpha + 1 > 0 \text{ and } \beta < \frac{\alpha \log 2}{\log (\alpha + 1)}, \tag{24}$$

$$3\beta - \alpha - 2 < 0. \tag{25}$$

2.9. From $(\alpha, \beta) \subset E$ we infer that the inequality

$$I(f, x, h, \alpha) \leqq A(f, x, h, \beta) \tag{26}$$

is satisfied, in particular, by every positive linear function $ax+b$. As observed in 2.3, this is equivalent to the fact that the function

$$\phi(u) = \log \frac{\left[\dfrac{1}{\alpha+1}\,\dfrac{u^{\alpha+1}-1}{u-1}\right]^{1/\alpha}}{\left(\dfrac{1+u^\beta}{2}\right)^{1/\beta}}$$

is $\leqq 0$ in $0<u<1$. On account of (22), we can use the lemmas developed in §1.

If either (23) or (24) holds, we have (see 1.7)

$$\phi(+0) > 0.$$

If (25) holds, then (see 1.6)

$$\operatorname{sgn} \phi(u) = -\operatorname{sgn}(3\beta - \alpha - 2) = +1 \quad \text{for } 1 - \epsilon < u < 1$$

where ϵ is some positive constant. Hence $\phi(u) \leqq 0$ is not satisfied. This contradicts (26), and the proof is complete.

3. Determination of the set $\overline{E}$

3.1. The set $\overline{E}$ has been defined in 0.13 as consisting of all pairs (α, β) such that the inequality

$$I(f, x, h, \alpha) \leqq A(f, x, h, \beta) \tag{27}$$

implies the convexity of the function $f(x)$. It is assumed that (27) holds for all values of x and h such that $x_1 < x-h < x+h < x_2$, and that $f(x)$ is positive and continuous in $x_1 < x < x_2$.

We shall now prove that $(\alpha, \beta) \subset \overline{E}$ if and only if $3\beta - \alpha - 2 \leqq 0$.

3.2. We shall need the following trivial remark: if $(\alpha, \beta) \subset \overline{E}$, and $\eta > 0$, then $(\alpha+\eta, \beta-\eta) \subset \overline{E}$.

Indeed, suppose that $f(x)$ is a positive and continuous function in $x_1 < x < x_2$ which satisfies the inequality

$$I(f, x, h, \alpha+\eta) \leqq A(f, x, h, \beta-\eta) \tag{28}$$

for all values of x and h such that $x_1 < x-h < x+h < x_2$. Since I and A are increasing functions of their last arguments, we have

$$I(f, x, h, \alpha) \leqq I(f, x, h, \alpha+\eta), \tag{29}$$

$$A(f, x, h, \beta-\eta) \leqq A(f, x, h, \beta). \tag{30}$$

(28), (29) and (30) yield

$$I(f, x, h, \alpha) \leqq A(f, x, h, \beta). \tag{31}$$

Since $(\alpha, \beta) \subset \bar{E}$, it follows from (31) that $f(x)$ is convex. Thus, starting from (28), we proved that $f(x)$ is convex. Hence (28) implies the convexity of $f(x)$, and consequently $(\alpha+\eta, \beta-\eta) \subset \bar{E}$.

3.3. We shall show first: if $3\beta-\alpha-2>0$, then (α, β) is not in $\bar{E}$. Suppose that

$$3\beta - \alpha - 2 > 0. \tag{32}$$

Let σ and η be two constants such that

$$\sigma > 1, \quad \eta > 0,$$

and put

$$\bar{\alpha} = \frac{\alpha + \eta}{\sigma}, \quad \bar{\beta} = \frac{\beta - \eta}{\sigma}\cdot \tag{33}$$

If σ is close enough to 1 and η close enough to zero and suitably chosen, then $(\bar{\alpha}, \bar{\beta})$ will satisfy the relations

$$\bar{\alpha} \neq 0, \ \bar{\alpha} + 1 \neq 0, \ \bar{\beta} \neq 0, \ 3\bar{\beta} - \bar{\alpha} - 2 \neq 0.$$

We set up the function

$$\bar{\phi}(u) = \log \frac{\left[\dfrac{1}{\bar{\alpha}+1}\,\dfrac{u^{\bar{\alpha}+1} - 1}{u - 1}\right]^{1/\bar{\alpha}}}{\left(\dfrac{1 + u^{\bar{\beta}}}{2}\right)^{1/\bar{\beta}}}, \quad 0 < u < 1.$$

On account of 1.6, we have an $\epsilon>0$, such that

$$\operatorname{sgn} \bar{\phi}(u) = -\operatorname{sgn}(3\bar{\beta} - \bar{\alpha} - 2) \text{ for } 1 - \epsilon < u < 1.$$

Hence, by (32),

$$\bar{\phi}(u) < 0 \text{ for } 1 - \epsilon < u < 1. \tag{34}$$

We define now a positive linear function l by the conditions

$$l(x_1) = 1 - \epsilon, \quad l(x_2) = 1, \ x_1 < x < x_2. \tag{35}$$

We assert that l satisfies the inequality

$$I(l, x, h, \bar{\alpha}) \leqq A(l, x, h, \bar{\beta}) \tag{36}$$

for all values of x and h such that $x_1<x-h<x+h<x_2$. If we put

$$u_0 = \frac{l(x - h)}{l(x + h)}, \tag{37}$$

then (see 2.3) this assertion is equivalent to the assertion

$$\bar{\phi}(u_0) \leqq 0. \tag{38}$$

But, from (35) and (37) and the linearity of $l(x)$ we have

$$1 - \epsilon < u_0 < 1. \tag{39}$$

Hence (38) and consequently (36) follow from (39) and (34).

Consider now the function

$$f(x) = l(x)^{1/\sigma}, \quad x_1 < x < x_2. \tag{40}$$

This function $f(x)$ is continuous and positive in $x_1<x<x_2$. If we raise (36) to the power $1/\sigma$, then by (33) and (40),

$$I(f, x, h, \alpha + \eta) \leqq A(f, x, h, \beta - \eta). \tag{40*}$$

On the other hand, we find from (40)

$$f''(x) = \frac{1}{\sigma}\left(\frac{1}{\sigma} - 1\right) l(x)^{(1-2\sigma)/\sigma}(l'(x))^2.$$

Since $\sigma>1$, we see that $f''(x)<0$ in $x_1<x<x_2$. Hence $f(x)$ is not convex in $x_1<x<x_2$.

To summarize, we have exhibited a function $f(x)$, which is continuous and positive in $x_1<x<x_2$ and which satisfies (40*) for all values of x and h such that $x_1<x-h<x+h<x_2$, and which is *not convex* in $x_1<x<x_2$. This shows that $(\alpha+\eta, \beta-\eta)$ is not in $\overline{E}$. From 3.2, it follows finally that (α, β) is not in $\overline{E}$ either.

3.4. We show next that if $3\beta-\alpha-2\leqq 0$, then $(\alpha, \beta) \subset \overline{E}$.† We first assume that the pair (α, β) also satisfies the conditions

$$3\beta - \alpha - 2 < 0, \alpha \neq 0, \alpha + 1 \neq 0, \beta \neq 0. \tag{41}$$

Let $f(x)$ be any positive continuous function in $x_1<x<x_2$ which satisfies the inequality

$$I(f, x, h, \alpha) \leqq A(f, x, h, \beta) \tag{42}$$

for all values of x and h such that $x_1<x-h<x+h<x_2$. We have to show that $f(x)$ is convex.

Suppose $f(x)$ is not convex. Then we have an x_0 and an h_0 such that $x_1<x_0-h_0<x_0+h_0<x_2$ and

$$f(x_0) > \frac{f(x_0 - h_0) + f(x_0 + h_0)}{2}.$$

† Cf. 0.15.

Clearly, we have then two numbers a_1, b_1 such that $x_0-h_0<a_1<b_1<x_0+h_0$ and such that the following is true: if $l_1(x)$ is the linear function defined by $l_1(a_1)=f(a_1)$, $l_1(b_1)=f(b_1)$, then $l_1<f$ in $a_1<x<b_1$.

3.5. By a simple reasoning whose details may be left to the reader,† we infer from this situation the existence of a number $\bar{x}$ and of two sequences a_n and b_n, such that the following hold:

I. $a_1<a_n<a_{n+1}<\bar{x}<b_{n+1}<b_n<b_1$.

II. $a_n\rightarrow\bar{x}$, $b_n\rightarrow\bar{x}$.

III. If l_n is the linear function defined by $l_n(a_n)=f(a_n)$, $l_n(b_n)=f(b_n)$, then

$$l_n \leqq f \text{ in } a_n \leqq x \leqq b_n.$$

IV. For $n=1$ we have, on account of 3.4, the stronger relation

$$l_1 < f \text{ in } a_1 < x < b_1.$$

V. All the lines $y=l_n(x)$ are parallel to each other.

3.6. From III in 3.5 we infer that

$$I(l_n, x_n, h_n, \alpha) \leqq I(f, x_n, h_n, \alpha), \tag{43}$$

where

$$x_n = \frac{a_n + b_n}{2}, \quad h_n = \frac{b_n - a_n}{2},$$

and also that

$$A(f, x_n, h_n, \beta) = A(l_n, x_n, h_n, \beta). \tag{44}$$

From (43), (44), (42) we see that

$$I(l_n, x_n, h_n, \alpha) \leqq A(l_n, x_n, h_n, \beta). \tag{45}$$

From IV in 3.5 it follows that for $n=1$ we have the stronger inequality

$$I(l_1, x_1, h_1, \alpha) < A(l_1, x_1, h_1, \beta). \tag{46}$$

Let us put

$$u_n = \frac{\min[l_n(a_n), l_n(b_n)]}{\max[l_n(a_n), l_n(b_n)]}.$$

Then

$$0 < u_n \leqq 1.$$

If $u_n=1$ for some n, then $l_n(x)$ is constant. Since all the lines $y=l_n(x)$ are

---

† Hint: Move the line $y=l_1(x)$ upward, parallel to itself, until it reaches a final position in which it still has some point in common with $y=f(x)$.

parallel to each other, we see that either $u_n=1$ for all values of n, or $0<u_n<1$ for all values of n.

3.7. Suppose first that $u_n=1$ for all values of n. Then, in particular, $l_1(x)$ reduces to some constant c, and (46) reduces to $c<c$. Thus this case is impossible.

3.8. We have therefore $0<u_n<1$ for all values of n. On account of II in 3.5 we have

$$u_n \to 1.$$

As observed in 2.3, the inequality (45) is equivalent to the fact that the function $\phi(u)$ of 1.5 satisfies

$$\phi(u_n) \leqq 0. \tag{47}$$

On account of 1.6, we have an $\epsilon>0$ such that

$$\operatorname{sgn} \phi(u) = -\operatorname{sgn}(3\beta-\alpha-2) \text{ for } 1-\epsilon<u<1.$$

Hence, with regard to (41),

$$\phi(u) > 0 \text{ for } 1-\epsilon<u<1. \tag{48}$$

Since $u_n \to 1$, (48) and (47) obviously contradict each other for large values of n. Thus the assumption that $f(x)$ is not convex is shown to lead to a contradiction.

3.9. We drop now the last three restrictions in (41) and we assume only that $3\beta-\alpha-2<0$. If $\eta>0$ is sufficiently small, then $\bar{\alpha}=\alpha-\eta$, $\bar{\beta}=\beta+\eta$ will satisfy all four conditions stated in (41). Hence, as proved above, $(\bar{\alpha}, \bar{\beta}) \subset \overline{E}$. From 3.2 it follows then that $(\alpha, \beta)=(\bar{\alpha}+\eta, \bar{\beta}-\eta)$ also belongs to $\overline{E}$.

3.10. It remains to show that if

$$3\beta-\alpha-2=0, \tag{49}$$

then $(\alpha, \beta) \subset \overline{E}$. Let $f(x)$ be any positive and continuous function in $x_1<x<x_2$ which satisfies

$$I(f, x, h, \alpha) \leqq A(f, x, h, \beta) \tag{50}$$

for all values of x and h such that $x_1<x-h<x+h<x_2$. Let γ be a constant such that

$$\gamma > 1,$$

and put

$$f(x)^\gamma = g(x).$$

Then it follows from (50) that $g(x)$ satisfies the inequality

$$I(g, x, h, \bar{\alpha}) \leqq A(g, x, h, \bar{\beta}), \tag{51}$$

where

$$\bar{\alpha} = \alpha/\gamma, \quad \bar{\beta} = \beta/\gamma.$$

From (49) we get

$$3\bar{\beta} - \bar{\alpha} - 2 = \frac{3\beta - \alpha - 2\gamma}{\gamma} = \frac{2(1-\gamma)}{\gamma} < 0. \tag{52}$$

But then, on account of 3.9, (51) implies that $g(x)$ is convex. That is to say: $f(x)^\gamma$ is convex, whenever $\gamma > 1$. Allowing $\gamma \to 1$ we conclude finally that $f(x)$ itself is also convex. Thus we see that (50) implies the convexity of $f(x)$, which proves that $(\alpha, \beta) \subset \bar{E}$.

4. Miscellaneous applications

4.1. In what precedes, we were concerned with convex functions. The lemmas developed in §1 cover however the case of concave functions also. Since the proofs result by obvious modifications of those we used for convex functions, we restrict ourselves to statements of results.

4.2. *Let us define the set E^* as consisting of all pairs (α, β) such that the following assertion is true: every function $f(x)$, which is positive, continuous and concave in $x_1 < x < x_2$, satisfies the inequality $I(f, x, h, \alpha) \geqq A(f, x, h, \beta)$ for all values of x and h such that $x_1 < x - h < x + h < x_2$.*

Theorem. *A pair (α, β) belongs to E^* if and only if one of the following five conditions is satisfied.*†

I. $\alpha \leqq -2$ *and* $\beta \leqq \dfrac{\alpha + 2}{3}$.

II. $-2 \leqq \alpha \leqq -1$ *and* $\beta \leqq 0$.

III. $-1 \leqq \alpha \leqq -\dfrac{1}{2}$ *and* $\beta \leqq \dfrac{\alpha \log 2}{\log(\alpha + 1)}$.

IV. $-\dfrac{1}{2} \leqq \alpha \leqq 1$ *and* $\beta \leqq \dfrac{\alpha + 2}{3}$.

V. $1 \leqq \alpha$ *and* $\beta \leqq \dfrac{\alpha \log 2}{\log(\alpha + 1)}$.

4.3. *Let us define the set $\bar{E}^*$ as consisting of all pairs (α, β) for which the following assertion is true: if a function $f(x)$, which is positive and continuous in $x_1 < x < x_2$, satisfies the inequality $I(f, x, h, \alpha) \geqq A(f, x, h, \beta)$ for all values of x and h such that $x_1 < x - h < x + h < x_2$, then $f(x)$ is concave in $x_1 < x < x_2$.*

† In terms of the functions $\psi_1(\alpha)$, $\psi_2(\alpha)$, $\psi_3(\alpha)$, used in the second footnote on p. 269, the set E^* may be described as consisting of all those points (α, β) for which

$$\beta \leqq \min\,[\psi_1(\alpha), \psi_2(\alpha), \psi_3(\alpha)].$$

THEOREM. *The pair* (α, β) *belongs to* $\overline{E}^*$ *if and only if* $3\beta-\alpha-2\geqq 0$.

4.4. We proceed now to present a few applications of the preceding results.

We shall say that the inequality

$$I(f, x, h, \alpha) \leqq A(f, x, h, \beta) \tag{53}$$

expresses a *characteristic property of positive and continuous convex functions*, if the following assertion is true: a function $f(x)$, positive and continuous in $x_1 < x < x_2$, is convex there if and only if it satisfies the inequality (53) for all values of x and h such that $x_1 < x-h < x+h < x_2$.

It is then clear that (53) expresses a characteristic property of positive and continuous convex functions if and only if the pair (α, β) is in both sets E and $\overline{E}$, defined in 0.13. From 0.13 we infer therefore the following

THEOREM. *The inequality*

$$I(f, x, h, \alpha) \leqq A(f, x, h, \beta)$$

expresses a characteristic property of positive and continuous convex functions if and only if (1) $3\beta-\alpha-2=0$ *and* (2) *either* $-2\leqq\alpha\leqq-\frac{1}{2}$ *or* $1\leqq\alpha<+\infty$.

4.5. In a similar fashion, from 4.2 and 4.3, we have the

THEOREM. *The inequality*

$$I(f, x, h, \alpha) \geqq A(f, x, h, \beta)$$

expresses a characteristic property of positive and continuous concave functions if and only if (1) $3\beta-\alpha-2=0$ *and* (2) *either* $\alpha\leqq-2$ *or* $-\frac{1}{2}\leqq\alpha\leqq 1$.

4.6. If we put $\beta=0$ in the theorem of 4.4, we see that

The inequality

$$\left[\frac{1}{2h}\int_{-h}^{h} f(x+\xi)^{-2}d\xi\right]^{-1/2} \leqq \textit{geometric mean of } f(x-h) \textit{ and } f(x+h)$$

is characteristic for positive and continuous convex functions. There does not exist any other characteristic inequality of the form

$$I(f, x, h, \alpha) \leqq (f(x-h)f(x+h))^{1/2}.$$

For $\beta=-1$, we obtain the following from 4.4:

There does not exist any inequality of the form

$$I(f, x, h, \alpha) \leqq \textit{harmonic mean of } f(x-h) \textit{ and } f(x+h),$$

which would be characteristic for positive and continuous convex functions.

We leave it to the reader to formulate the corresponding theorems for concave functions, on the basis of 4.5.

4.7. The results developed in this paper imply an infinity of *sharp inequalities for convex and for concave functions*. We wish to illustrate this point on the following special example. *Let us ask for the best inequality of the form*

$$I(f, x, h, 2) \leqq A(f, x, h, \beta), \tag{54}$$

which is satisfied by all positive and continuous convex functions. Since the right-hand side of (54) is an increasing function of β, our problem requires the determination of the smallest value of β such that (54) holds for every continuous and positive convex function. In other words, we have to determine the smallest β such that the pair $(2, \beta)$ is in the set E defined in 0.13. It follows from 0.13 that this smallest value is 4/3. Hence:

The inequality

$$\left[\frac{1}{2h}\int_{-h}^{h} f(x+\xi)^2 d\xi\right]^{1/2} \leqq \left[\frac{f(x-h)^{4/3}+f(x+h)^{4/3}}{2}\right]^{3/4} \tag{55}$$

is a sharp inequality for positive and continuous convex functions, in the following sense. Every positive and continuous convex function satisfies (55), *but for every* $\epsilon > 0$ *there exist positive and continuous convex functions which do not satisfy the inequality*

$$\left[\frac{1}{2h}\int_{-h}^{h} f(x+\xi)^2 d\xi\right]^{1/2} \leqq \left[\frac{f(x-h)^{4/3-\epsilon}+f(x+h)^{4/3-\epsilon}}{2}\right]^{1/(4/3-\epsilon)}.$$

Clearly, our results permit us to determine, for every fixed α and for every fixed β, the sharp inequality of the form $I(f, x, h, \alpha) \leqq A(f, x, h, \beta)$ for convex functions, and to answer the corresponding questions for concave functions.

4.8. As a last application, we shall present and discuss a theorem which gives a complete answer to the question raised in 0.9 concerning the classes C_γ and C_δ^*.

THEOREM. *The relation* $C_\gamma \equiv C_\delta^*$ *holds if and only if* (1) $2\gamma + \delta - 3 = 0$ *and* (2) *either* $\gamma \leqq 1$ *or* $\gamma \geqq 2$.

This theorem is an immediate consequence of our results concerning the sets E, $\overline{E}$, E^*, $\overline{E}^*$. We shall reproduce the reasoning in the case $\gamma > 0$. The verification of the theorem for the cases $\gamma < 0$ and $\gamma = 0$ will be left to the reader.

4.9. Suppose then that $\gamma > 0$ and let us ask for all those values of δ, if any, for which $C_\gamma \equiv C_\delta^*$.

If $f(x)$ is positive and continuous in $x_1<x<x_2$, then $f\subset C_\gamma$ means that $g=f^\gamma$ is convex. On the other hand, $f\subset C_\delta^*$ means that $I(f,x,h,\delta)\leqq A(f,x,h,1)$ for all values of x and h such that $x_1<x-h<x+h<x_2$. This inequality can be written, in terms of the function $g=f^\gamma$, as

$$I(g, x, h, \delta/\gamma) \leqq A(g, x, h, 1/\gamma). \tag{56}$$

Hence, $C_\gamma\equiv C_\delta^*$ is equivalent to the fact that (56) is a necessary and sufficient condition for the convexity of g. In other words, $C_\gamma\equiv C_\delta^*$ is equivalent to the fact that (56) expresses a characteristic property of positive and continuous convex functions. Putting

$$\alpha = \frac{\delta}{\gamma}, \quad \beta = \frac{1}{\gamma},$$

the theorem of 4.4 yields the following necessary and sufficient conditions:

$$3\beta - \alpha - 2 = \frac{3 - \delta - 2\gamma}{\gamma} = 0, \text{ and} \tag{i}$$

$$\text{either } -2 \leqq \alpha = \frac{\delta}{\gamma} \leqq -\frac{1}{2} \text{ or } 1 \leqq \alpha = \frac{\delta}{\gamma}. \tag{ii}$$

These conditions are equivalent to the following set:

$$\delta = 3 - 2\gamma, \text{ and} \tag{i$'$}$$

$$\text{either } -2\gamma \leqq 3 - 2\gamma \leqq -\tfrac{1}{2}\gamma \text{ or } \gamma \leqq 3 - 2\gamma, \tag{ii$'$}$$

which is finally equivalent to the set

$$2\gamma + \delta - 3 = 0, \text{ and} \tag{i$''$}$$

$$\text{either } \gamma \geqq 2 \text{ or } \gamma \leqq 1. \tag{ii$''$}$$

These are, however, exactly the conditions stated in 4.8.

4.10. It is interesting to compare the theorem in 4.8 with the remark made in 0.9. According to 0.9, we have

$$C_\delta^* \subset C_\gamma \text{ for } 2\gamma + \delta - 3 = 0. \tag{57}$$

For what values of γ *does the converse hold*? In other words: for what values of γ is it true that

$$C_\gamma \subset C_\delta^* \text{ for } 2\gamma + \delta - 3 = 0? \tag{58}$$

With regard to (57), if (58) holds for a certain γ, then $C_\gamma\equiv C_\delta^*$, where γ and

δ are related by the equation $2\gamma+\delta-3=0$. According to 4.8, we have this situation if and only if either $\gamma \leqq 1$ or $\gamma \geqq 2$. Summing up:† *we have*

$$C_\delta^* \subset C_\gamma \textit{ for } 2\gamma + \delta - 3 = 0$$

without any further restriction on γ. *The converse, namely the relation*

$$C_\gamma \subset C_\delta^* \textit{ for } 2\gamma + \delta - 3 = 0,$$

holds however if and only if $\gamma \leqq 1$ *or* $\gamma \geqq 2$. *For* $1 < \gamma < 2$ *the converse is false.*

† Cf. 0.9 for the origin of this theorem.

Ohio State University,
Columbus, Ohio

SYSTEMS OF ALGEBRAIC MIXED DIFFERENCE EQUATIONS*

BY
FRITZ HERZOG

In his algebraic theory of differential equations, J. F. Ritt† has developed a decomposition theory for systems of algebraic differential equations by introducing the idea of irreducible systems and proving that every system is equivalent to one and essentially only one finite set of irreducible systems. The analogous theorem for algebraic difference equations was given by Ritt and Doob.‡ The purpose of the present paper is to derive a decomposition theorem for algebraic mixed difference equations; i.e., equations which contain algebraically one or more unknown functions $y(x)$, their "transforms" $y(x)$, $y(x+1)$, $y(x+2)$, $\cdots$, and the derivatives of those transforms.

1. Fields

Let $\mathfrak{A}$ be an open region in the plane of the complex variable x which has the property that it contains the point $x+1$ whenever x lies in $\mathfrak{A}$. Let $\mathfrak{F}$ be a set of functions, meromorphic in $\mathfrak{A}$. If the following four conditions are satisfied $\mathfrak{F}$ will be called a *field*:

(a) $\mathfrak{F}$ contains at least one function which is not identically zero.
(b) If $f(x)$ and $g(x)$ are in $\mathfrak{F}$, then $f\pm g$, $f\cdot g$ and $f/g(g\neq 0)$ are in $\mathfrak{F}$.
(c) If $f(x)$ is in $\mathfrak{F}$ then also $f(x+1)$ is in $\mathfrak{F}$.
(d) If $f(x)$ is in $\mathfrak{F}$ then also the derivative of $f(x)$ is in $\mathfrak{F}$.

2. Forms

We introduce a finite number of letters $y_1, y_2, \cdots, y_n$, which will represent unknown functions $y_1(x), y_2(x), \cdots, y_n(x)$ and will be called *unknowns*. With every y_i we associate the symbols

$$y_{ij} = y_i(x+j) \qquad (y_{i0} = y_i),$$

$$y_{ijk} = \frac{d^k y_{ij}}{dx^k} = \frac{d^k y_i(x+j)}{dx^k} \qquad (y_{ij0} = y_{ij},\ y_{i00} = y_{i0} = y_i)$$

---

* Presented to the Society, October 27, 1934; received by the editors August 9, 1934.

† Ritt, *Differential Equations from the Algebraic Standpoint*, Colloquium Publications of the American Mathematical Society, vol. 14, 1932.

‡ Ritt and Doob, *Systems of algebraic difference equations*, American Journal of Mathematics, vol. 55 (1933), pp. 505–514.

where j and k are non-negative integers. $y_{ij}=y_i(x+j)$ will be called the *jth transform of the unknown* y_i.

A polynomial in the y_{ijk} the coefficients of which are functions, meromorphic in an open region $\mathfrak{A}$, will be called a *form* in the unknowns $y_1, \cdots, y_n$. Throughout our work we shall assume that the forms with which we deal belong to a fixed field $\mathfrak{F}$. Forms will be denoted by capital italic letters.

By the *class* of a form F, if F involves one or more y_{ijk} effectively, we mean the highest index p, such that some of the transforms y_{pj} of y_p or some of their derivatives y_{pjk} are present in F. By forms of class zero, we mean forms free of the y_{ijk}, i.e., functions $f(x)$ of the field $\mathfrak{F}$.

By the *Δ-order* of F in y_h, if F involves some y_{hjk} ($j\geqq 0$, $k\geqq 0$), we mean the order of the highest transform y_{hq} of y_h such that y_{hq} or some of its derivatives y_{hqk} are present in F.

By the *d-order* of F in y_{hl}, if F involves some y_{hlk} ($k\geqq 0$), we mean the order of the highest derivative of y_{hl} which is present in F.

From what precedes, it follows that we can associate four integers with each form in $y_1, \cdots, y_n$, not of class zero, viz.

(i) the class $p(p>0)$,

(ii) the Δ-order q in y_p ($q\geqq 0$),

(iii) the d-order r in y_{pq} ($r\geqq 0$),

(iv) the degree s in y_{pqr} (in the ordinary algebraic sense; $s>0$).

These four numbers, in the succession described above,

$$p,\ q,\ r,\ s,$$

which will be called simply the *quadruplet* of the form, define the *rank* of the forms in the following sense.

If A and B are forms in the same unknowns $y_1, y_2, \cdots, y_n$, neither of class zero, if

$$a_1,\ a_2,\ a_3,\ a_4$$

and

$$b_1,\ b_2,\ b_3,\ b_4$$

are respectively the quadruplets of A and B, then A and B are said to be *of the same rank* if $a_i=b_i$ for $i=1, 2, 3, 4$. B will be said to be *of lower rank* than A if, for the smallest index i for which a_i and b_i are different, b_i is less than a_i.

As for the forms of class zero, all of them are to be of the same rank and each of them of lower rank than a form of non-zero class.

Obviously, if A is higher than B and B higher than C, then A is higher than C.

The following lemma can be proved very easily.*

LEMMA. *Every set of forms in $y_1, y_2, \cdots, y_n$ contains at least one form, the rank of which is not higher than that of any other form of the set.*

3. Ascending sets

Let F be a form, not of class zero, with the quadruplet p, q, r, s $(p>0)$. A form H is said to be *reduced with respect to the form* F if for every $k \geqq q$ either

(a) H does not involve y_{pk}, or

(b) the d-order of H in y_{pk} is less than r, or

(c) the d-order of H in y_{pk} equals r but the degree of H in y_{pkr} is less than s.†

In the following we shall sometimes say "H is of lower rank in y_{pk} than r, s" if H has any one of the properties (a), (b), or (c) with respect to y_{pk}.

A set of forms

$$A_1, A_2, \cdots, A_m$$

is called an *ascending set* if either

(a) it consists of only one form A_1, not identically zero, or

(b) it contains more than one form, A_1 is of class greater than zero, and for $1 \leqq i < j \leqq m$, A_j is reduced with respect to and of higher rank than A_i.

Let

$$(A) \qquad A_1, A_2, \cdots, A_k,$$

$$(B) \qquad B_1, B_2, \cdots, B_l$$

be two ascending sets. (A) will be called *of higher rank* than (B) if either

(a) there exists a j, exceeding neither k nor l, such that A_i and B_i are of the same rank for $i=1, 2, \cdots, j-1$, but A_j is higher than B_j,‡

or

(b) l is greater than k and A_i and B_i have the same rank for $i=1, 2, \cdots, k$. If $l=k$ and A_i and B_i have the same rank for $i=1, 2, \cdots, k$, the ascending sets (A) and (B) are to be *of the same rank*.

Evidently, if (A), (B), (C) are ascending sets, (A) higher than (B), and (B) higher than (C), then (A) is also higher than (C).

We shall now prove the following

* Cf. Ritt, loc. cit., p. 3, and Ritt-Doob, loc. cit., p. 506.

† It may be that for different k's we have different cases.

‡ In the case $j=1$, that means simply that A_1 is higher than B_1.

LEMMA. *If Φ is a system of ascending sets there exists in Φ at least one ascending set of lowest rank.*

Consider the first forms of all ascending sets of Φ. According to the lemma of §2, there are among them forms of lowest rank. Let Φ_1 consist of all such ascending sets of Φ that begin with a form of that lowest rank. If all ascending sets of Φ_1 have only one form, then each ascending set of Φ_1 constitutes an ascending set of lowest rank in Φ.* If not, consider the ascending sets of Φ_1 which have at least two forms and choose those from among them whose second form is of lowest rank. Let the subsystem of Φ_1 obtained in this way be called Φ_2. If all ascending sets of Φ_2 have exactly two forms each of them constitutes an ascending set of lowest rank in Φ. If not, we continue in this fashion. If we can show that, after a finite number of steps, we reach a system Φ_m all ascending sets of which have exactly m forms, then each ascending set of Φ_m constitutes an ascending set of lowest rank in Φ and the lemma will be proved.

In order to show that we must reach such a system Φ_m, we consider any system Φ_i which we meet in the above process. We notice that all ith forms of the ascending sets of Φ_i are of the same rank. Let A_i be one of these forms and p_i, q_i, r_i, s_i be the quadruplet of A_i.† Consider now any two consecutive systems Φ_j and Φ_{j+1}. Since A_{j+1} is of higher rank than and reduced with respect to A_j we have only the following possibilities:

(a) $$p_{j+1} > p_j,$$

(b) $$p_{j+1} = p_j;\ q_{j+1} > q_j.$$

In case (b), since A_{j+1} is of lower rank in $y_{p_j q_{j+1}}$ than r_j, s_j, either

(b′) $$r_{j+1} < r_j$$

or

(b″) $$r_{j+1} = r_j;\ s_{j+1} < s_j.$$

The inequalities in (b′) and (b″) now show that the case (b) can occur only for a finite number of consecutive systems Φ_i. Therefore, after a finite number of steps the case (a) must hold, i.e., the class p_i must increase. But since p_i can assume only the values $1, 2, \cdots, n$ the process cannot be continued ad infinitum.

---

* This includes the case in which at least one set of Φ consists of a non-zero form of class zero. Every such ascending set is, of course, of lowest rank.

† We omit the trivial case, mentioned above, in which A_1 is of class zero; then $m=1$.

4. Basic sets

If Σ is a set of forms not all of which vanish identically, Σ contains ascending sets (e.g., every form $F\not\equiv 0$ constitutes an ascending set). Among them, according to the lemma of §3, there are ascending sets of lowest rank. Each such ascending set will be called a *basic set* of Σ.

If a form F is reduced with respect to every form of an ascending set

$$A_1, A_2, \cdots, A_m \tag{1}$$

whose first form A_1 is not of class zero, then F is said to be ***reduced with respect to the ascending set*** (1).

We prove the following lemma:

Lemma 1. *If Σ is a system of forms,* (1) *one of its basic sets, with A_1 not of class zero, and if F is a non-zero form and reduced with respect to* (1), *then $\Sigma+F$ has a basic set lower than* (1).

F cannot be of the same rank as any of the forms in (1). If F is lower than A_1 the ascending set F is lower than (1). If not, let j be the highest index such that A_j is lower than F $(1\leqq j\leqq m)$. Then $A_1, \cdots, A_j, F$ is lower than (1).

In exactly the same way one can prove

Lemma 2. *If Σ is a system of forms,* (1) *one of its basic sets, with A_1 not of class zero, then Σ contains no non-zero form which is reduced with respect to* (1).

5. Reduction of forms

If A is any form in $y_1, y_2, \cdots, y_n$, we shall mean by the ***kth transform of the form A*** the form which is obtained on replacing x by $x+k$ in the coefficients of A and in the unknowns, their transforms and their derivatives. The latter means that y_{ijl} has to be replaced by $y_{i,j+k,l}$. The kth transform will be denoted by $A^{(k)}[A^{(0)}=A]$.

Similarly, by the ***mth derivative of a form B*** we mean the mth total derivative of B with respect to x, the y_{ijl} being regarded as functions of x. The mth derivative of the kth transform of A will be denoted by $A^{(k,m)}[A^{(k,0)}=A^{(k)};$ $A^{(0,0)}=A^{(0)}=A]$.

Evidently, $A^{(k,m)}$ can be interpreted also as the kth transform of the mth derivative of A. In any case, the first superscript of a form indicates the number of transformations, the second one the number of differentiations.

The properties (c) and (d) in the definition of a field (§1) imply that the transforms and their derivatives of a form A have coefficients in $\mathfrak{F}$.

Let A be a form, not of class zero, with the quadruplet p, q, r, s. Let A be written as a polynomial in y_{pqr} with coefficients which are forms in the y_{ijk} lower than y_{pqr}. Then the derivative of this polynomial, i.e., $\partial A/\partial y_{pqr}$,

will be called the *separant* of A, and its highest coefficient, i.e., the coefficient of the term $(y_{pqr})^s$ in A, will be called the *initial* of A.

For the proof of the following reduction lemma we need a few facts about the quadruplet, the separant and the initial of the transforms and their derivatives of a form A, not of class zero. Let p, q, r, s be the quadruplet of A, S its separant, and I its initial. Then the following table gives the facts that we will need in the proof. They will be seen to follow easily from the above definitions.

(2)

Form	$A^{(k)}(k \geqq 0)$	$A^{(k,m)}(k \geqq 0, m > 0)$
Quadruplet	$p, q+k, r, s$	$p, q+k, r+m, 1$
Separant	$S^{(k)}$	$S^{(k)}$
Initial	$I^{(k)}$	$S^{(k)}$

We finally notice that the separant and the initial of any form are of lower rank than the form itself.

We are now ready to prove the

REDUCTION LEMMA. *Let*

$$A_1, A_2, \cdots, A_m \tag{3}$$

be an ascending set, with A_1 of class greater than zero, and let the separant and initial of the form A_i be S_i and I_i respectively $(i=1, \cdots, m)$. If G is an arbitrary form there exists a power product T of the S_i, the I_i, and their transforms with non-negative integral exponents, such that by subtracting from $T\cdot G$ a suitable linear combination of the A_i, their transforms $A_i^{(k)}$ and the derivatives $A_i^{(k,m)}$ of their transforms, with forms for coefficients, a form

$$R = T\cdot G - \sum_{i,k,m} H_{ikm}A_i^{(k,m)} \tag{4}$$

is obtained, which is reduced with respect to the ascending set (3).

It will suffice to consider the case in which G is not reduced with respect to (3).

Let j be the highest index such that G is not reduced with respect to A_j. Let p, q, r, s be the quadruplet of A_j. Since G is not reduced with respect to A_j there must exist at least one transform y_{ph} of y_p, with $h \geqq q$, such that the rank of G in y_{ph} is not lower than r, s. We may assume that y_{ph} is the highest such transform. Let ρ be the d-order of G in y_{ph} and let σ be the degree of G in $y_{ph\rho}$. Then either

(a) $$\rho > r$$

or

(b) $$\rho = r;\ \sigma \geqq s.$$

We treat first the case (a).

We take the form $A_j^{(h-q,\rho-r)}$ where, according to our assumptions, $h-q\geqq 0$, $\rho-r>0$. The quadruplet of $A_j^{(h-q,\rho-r)}$ is $p, h, \rho, 1$, its initial $S_j^{(h-q)}$ (cf. table (2)). Dividing G by $A_j^{(h-q,\rho-r)}$ we get, with a suitable integer $v\geqq 0$,

$$[S_j^{(h-q)}]^v\cdot G = C\cdot A_j^{(h-q,\rho-r)} + D. \tag{5}$$

For the sake of uniqueness we take v as small as possible.

D is either identically zero or of lower degree in $y_{ph\rho}$ than $A_j^{(h-q,\rho-r)}$, i.e., D does not contain $y_{ph\rho}$. Since higher derivatives of y_{ph} than the ρth do not appear in G and $A_j^{(h-q,\rho-r)}$, D cannot contain such derivatives either, so that the d-order of D in y_{ph} is less than ρ.

Furthermore, if y_{kl} is any transform higher than y_{ph}, D is not of higher rank in y_{kl} than G. Suppose that were not true for some y_{kl}. Notice that $A_j^{(h-q,\rho-r)}$ and $S_j^{(h-q)}$ are free of y_{kl}. From (5) we conclude that C would have to contain y_{kl} in the same rank, say μ,ν, as D. Hence $C\cdot A_j^{(h-q,\rho-r)}$ would contain a term involving the product $(y_{kl\mu})^\nu\cdot y_{ph\rho}$, which could be balanced neither by D nor by the first member of (5). This proves our statement.

We conclude first that D is of lower rank than r, s in each higher transform of y_p than the hth one, and secondly, that (in case $j<m$) D is reduced with respect to $A_{j+1}, A_{j+2}, \cdots, A_m$.

If now D involves y_{ph} and the d-order of D in y_{ph}, which was less than ρ, is still greater than r, i.e., if the case (a) still holds, we treat D in the same way as we did G, and, after a finite number (at most $\rho-r$) of such divisions, we reach a form D_1, which is either free of y_{ph} (and its derivatives) or else of d-order less than or at most equal to r in y_{ph} but is, as D was, of lower rank than r, s in $y_{p,h+1}, y_{p,h+2}, \cdots$, and reduced with respect to A_{j+1}, $A_{j+2}, \cdots, A_m$.

If D_1 involves y_{ph} and is of d-order exactly r in y_{ph} and of degree $\sigma\geqq s$ in y_{phr} (case (b)), then we take $A_j^{(h-q)}$ which has, according to table (2), the quadruplet p, h, r, s and the initial $I_j^{(h-q)}$. By a division we get, with some integer $w\geqq 0$ which we shall take as small as possible,

$$[I_j^{(h-q)}]^w\cdot D_1 = B\cdot A_j^{(h-q)} + D_2, \tag{6}$$

where D_2 is either free of y_{phr} or of degree less than s in y_{phr}.

Similarly, as we concluded for D from (5), we now conclude for D_2 from (6) that D_2 does not contain any higher derivative of y_{ph} than the rth one, so that D_2 is of lower rank in y_{ph} than r, s and, furthermore, that D_2 is of lower rank than D_1 in each y_{kl} which is higher than y_{ph}. This latter fact implies, as before, that D_2 is of lower rank than r, s also in $y_{p,h+1}, y_{p,h+2}, \cdots$, and that it is reduced with respect to $A_{j+1}, A_{j+2}, \cdots, A_m$.

Using all the divisions of the type (5) and (6) made so far, we have a relation between G and D_2 of the following type:

$$[S_j^{(h-q)}]^u \cdot [I_j^{(h-q)}]^w \cdot G - \sum_{t=0}^{p-r} E_t \cdot A_j^{(h-q,t)} = D_2$$

with non-negative integral u and w.

If now D_2 is not yet reduced with respect to A_j we must have, from what has been said about D_2, that there exists at least one transform $y_{ph'}$ of y_p, with $q \leqq h' < h$, such that D_2 is not of lower rank in $y_{ph'}$ than r, s. We then apply the same method as above to the highest such transform $y_{ph'}$ and reduce D_2 with respect to that transform. If necessary, we continue in this fashion until, after a finite number (at most $h-q+1$) of such steps, we have reached a form D_3, which is of lower rank than r, s in $y_{pq}, y_{p,q+1}, \cdots$, hence reduced with respect to A_j, but also, as D_2 before, with respect to $A_{j+1}, \cdots, A_m$.

As relation between G and D_3, we have

$$\prod_{\lambda=0}^{h-q} [S_j^{(\lambda)}]^{u_\lambda} [I_j^{(\lambda)}]^{w_\lambda} \cdot G - \sum_{k=0}^{h-q} \sum_m F_{km} A_j^{(k,m)} = D_3,$$

with non-negative integral exponents u_λ and w_λ.

If D_3 is not yet reduced with respect to (3) let i be the highest index such that D_3 is not reduced with respect to A_i. Then i must be less than j. We then apply the above method in order to reduce D_3 with respect to A_i and get a form D_4, reduced with respect to A_i and also to $A_{i+1}, \cdots, A_m$. After a finite number (at most j) of such steps we finally reach a form R which is reduced to the ascending set (3) and related to G by the equation (4). We have thus proved the reduction lemma.

The form R will be called in the sequel the *remainder of the form G with respect to the ascending set* (3).

6. Solutions

We shall define what will be meant by a solution of a system of algebraic

mixed difference equations. Our definition will be broad enough so as to admit also multiple-valued functions as solutions.*

Let $\mathfrak{A}$ be an open region of the type mentioned in §1. Let Σ be a system of forms in the unknowns $y_1, y_2, \cdots, y_n$, with coefficients meromorphic in $\mathfrak{A}$ and belonging to the field $\mathfrak{F}$ (cf. §2).

Let $\phi(t)$ be a continuous complex function, defined for real $t \geqq 0$. such that all values $\phi(t)$ lie in $\mathfrak{A}$ and $\phi(t)$ satisfies the equation

$$\phi(t+1) - \phi(t) = 1. \tag{7}$$

The function $x = \phi(t)$ defines a continuous path $\mathcal{L}$, lying entirely in $\mathfrak{A}$ and containing $x+1$ if x lies on it. By a function *analytic on* $\mathcal{L}$, we shall understand a function $f(x)$ which is analytic at the point $x = \phi(0)$ and which can be continued analytically along $\mathcal{L}$, as one moves on it with increasing t. For the points x on $\mathcal{L}$, $f(x) = f(\phi(t)) = \psi(t)$ becomes thus a unique function of t. Using (7), we can and shall define $f(x+1)$ by

$$f(x+1) = \psi(t+1).$$

It is obvious that $f(x+1)$ is also analytic on $\mathcal{L}$, and the same holds for all transforms and their derivatives of $f(x)$.

By a *solution* of the system Σ, we mean the entity composed of a path $\mathcal{L}$, as described above, and n functions $y_1(x), y_2(x), \cdots, y_n(x)$ analytic on $\mathcal{L}$, such that the forms of Σ vanish identically on $\mathcal{L}$ if the unknowns y_i in them are replaced by the functions $y_i(x)$.

The totality of solutions of a system Σ will be called the *manifold* of Σ.

If Σ_1 and Σ_2 are systems of forms such that every solution of Σ_1 is a solution of Σ_2, we shall say "Σ_2 *holds* Σ_1."† If two systems hold one another, i.e., if their manifolds are identical, they will be said to be *equivalent*.

We notice the following easy consequence of our definitions:

If $f(x)$ is analytic on $\mathcal{L}$, then $f(x+1) \equiv 0$ on $\mathcal{L}$ if and only if $f(x) \equiv 0$ on $\mathcal{L}$. Therefore, if F is any form, F has the same manifold as each of its transforms $F^{(k)}(k = 1, 2, \cdots)$.

7. Completeness

In this section we shall prove the following

Theorem. *Every infinite system of forms is equivalent to one of its finite subsystems.*

We call a system which is equivalent to one of its finite subsystems a *complete* system. Systems which are not complete will be called *incomplete*.

* Cf. Ritt-Doob, loc. cit., p. 506, footnote.

† A system devoid of solutions is considered as being held by every system.

We have to show that the assumption of the existence of incomplete systems leads to a contradiction.

We first prove the following

LEMMA. *Let Σ be an incomplete system. Let*

$$F_1, F_2, \cdots, F_m$$

be forms such that, by multiplying each form in Σ by some power product of the F_i and their transforms $F_i^{(k)}$, a system Λ is obtained which is complete. Then at least one of the systems $\Sigma + F_i$ $(i=1, 2, \cdots, m)$ is incomplete.*

Suppose each system $\Sigma + F_i$ is complete. Let Φ_i be a finite subsystem of $\Sigma + F_i$ which is held by $\Sigma + F_i$ $(i=1, 2, \cdots, m)$. Since Φ_i can be replaced by any finite subsystem of $\Sigma + F_i$ containing Φ_i in itself, we may assume that Φ_i is of the type

$$G_1, G_2, \cdots, G_l, F_i \tag{8}$$

(for $i=1, 2, \cdots, m$), with the set

$$G_1, G_2, \cdots, G_l \tag{9}$$

independent of i and the forms of (9) in Σ. We may, similarly, assume that the forms in Λ, obtained from the forms in (9) by the above described multiplication, form a system which is equivalent to Λ.

Since Σ is incomplete there is a form L in Σ which does not hold (9). But the product of L and some power product of the F_i and their transforms holds (9). Since F_i has the same manifold as its transforms (§6) we conclude that $F_1 \cdots F_m L$ holds (9). Since L does not hold (9) there are solutions of (9) for which L does not vanish, so that at least one of the F_i must vanish. Hence, for at least one i, L does not hold (8). This contradiction proves our lemma.

We shall now prove the completeness of all systems.

Suppose there are incomplete systems. Consider their totality and choose for each of them a basic set (§4). According to the lemma of §3, there are incomplete systems whose basic sets are not higher than those of any incomplete system. Let Σ be one of these systems, let

$$A_1, A_2, \cdots, A_m \tag{10}$$

be a basic set of Σ, and let S_i and I_i, respectively, be the separant and the initial of $A_i (i=1, 2, \cdots, m)$.

Then A_1 is of class greater than zero, for otherwise Σ would have no solutions and would be equivalent to A_1. Hence, it would be complete.

* The power product may be different for different forms in Σ.

We can then apply the reduction lemma (§5) to each form G of Σ not in (10), and form the remainder R of G with respect to the ascending set (10). Let T be the power product of the S_i, the I_i and their transforms, used in the reduction of G (cf. (4)). Let Ω be the system composed of the forms in (10) and the remainders R of all forms G of Σ not in (10). Let Λ be the system composed of the forms in (10) and the products $T \cdot G$ for all forms G of Σ not in (10).

Now the system Ω is complete. If not, Ω would contain non-zero forms, different from the forms in (10). Since such forms would be reduced with respect to (10), (10) could not be a basic set of Ω, according to Lemma 2 of §4. The basic sets of Ω then would have to be lower than (10), so that Ω would be an incomplete system with a basic set lower than (10). Thus Ω is complete.

Since every form in Λ which is not in (10) differs from the corresponding form in Ω by a linear combination of the A_i, their transforms and their derivatives, Λ and Ω are equivalent and Λ is also complete.

Applying the lemma of this section, we see that at least one of the $2m$ systems $\Sigma+S_i$ and $\Sigma+I_i$ is incomplete. But every S_i and I_i is not identically zero and reduced with respect to (10). Consequently, according to Lemma 1 of §4, each of the systems $\Sigma+S_i$ and $\Sigma+I_i$ has a lower basic set than (10). This contradiction proves that incomplete systems cannot exist.

8. Decomposition theorem

A system Σ of forms is said to be *reducible* if there exist two forms G and H, such that neither G nor H holds Σ but $G \cdot H$ does. Otherwise Σ is called *irreducible*.

A system Σ will be said to be *equivalent to the set of systems* $\Sigma_1, \Sigma_2, \cdots, \Sigma_r$ if every solution of Σ is a solution of at least one of the systems Σ_i and if every Σ_i is held by Σ. In other words, Σ is equivalent to the set $\Sigma_1, \Sigma_2, \cdots, \Sigma_r$ if the manifold of Σ is the sum of the manifolds of the Σ_i $(i=1, 2, \cdots, r)$.

We are now ready to prove the

Decomposition Theorem. *Every system Σ of forms in $y_1, y_2, \cdots, y_n$ is equivalent to a finite set of irreducible systems*

$$\Sigma_1, \Sigma_2, \cdots, \Sigma_r. \tag{11}$$

When those among the systems in (11) which are held by some other are suppressed, the set of the remaining systems, say

$$\Sigma_1, \Sigma_2, \cdots, \Sigma_s, \tag{12}$$

will still be equivalent to Σ but such that no system in (12) is held by any

other. Such a decomposition of Σ will be called a decomposition into *essential* irreducible systems.

The decomposition into essential irreducible systems is unique in the following sense: if (12) *and*

$$\Omega_1, \Omega_2, \cdots, \Omega_t \tag{13}$$

are two decompositions of Σ *into essential irreducible systems, then* $t=s$ *and, after a suitable permutation of the indices in* (13), Ω_i *and* Σ_i *are equivalent for* $i=1, 2, \cdots, s$.

We prove first that every system can be decomposed into a finite set of irreducible systems. Suppose that were not the case for some system Σ. Then Σ must be reducible. Let G_1, H_1 be such that $G_1 \cdot H_1$ holds Σ but neither G_1 nor H_1 does. Then Σ is equivalent to the set of systems $\Sigma+G_1$, $\Sigma+H_1$.

At least one of these two systems cannot be equivalent to a finite set of irreducible systems. Suppose $\Sigma+G_1$ is not. Then, similarly, we find a form G_2 such that

(a) G_2 does not hold $\Sigma+G_1$,

(b) $\Sigma+G_1+G_2$ is not equivalent to a finite set of irreducible systems.

Continuing, we find an infinite sequence of forms

$$G_1, G_2, \cdots, G_k, \cdots \tag{14}$$

such that, for $i=1, 2, \cdots$, G_i does not hold the system $\Sigma+G_1+G_2+\cdots+G_{i-1}$.

Let Λ be the system composed of the forms in Σ and all the forms in (14). According to the theorem of §7, Λ is equivalent to some finite subsystem Φ of Λ. Let j be such that no G_i in (14) with a higher index than j belongs to Φ. Then Λ is equivalent also to its subsystem,

$$\Sigma + G_1 + G_2 + \cdots + G_j. \tag{15}$$

But that is impossible, since G_{j+1} does not hold (15). This contradiction shows the existence of a decomposition.

We have now to show the uniqueness of the decomposition.

Let us consider the two decompositions (12) and (13) of Σ into essential irreducible systems. We shall show first that Σ_1 is held by some Ω_i. Suppose none of the systems in (13) holds Σ_1. Let B_i be a form in Ω_i such that B_i does not hold Σ_1 ($i=1, 2, \cdots, t$). Then the form $B_1 \cdots B_t$ would hold each Ω_i, hence Σ and also Σ_1. This is impossible since Σ_1 is irreducible.

Let Ω_1 hold Σ_1. Similarly, Ω_1 must be held by some Σ_i. Since Σ_1 is held by no Σ_i with $i>1$, Ω_1 is held by Σ_1. Consequently Ω_1 and Σ_1 are equivalent. Continuing in this way, we get the result that each Σ_i in (12) is equivalent to

some Ω_j in (13) and vice versa, so that t and s must be equal. Thus the decomposition is unique in the sense described above.

9. Examples

1. Let us consider the system Σ consisting of the following two equations in one unknown y:

$$y(x+1)\cdot y'(x)^2 + \pi^2\cdot e^{2\pi i x}\cdot y(x) = 0, \tag{16}$$

$$y(x+2) - y(x) = 0. \tag{17}$$

Let $\mathfrak{F}$ consist of all rational functions in $e^{2\pi i x}$ with complex coefficients.

Multiplying (17) by $y(x+1)$, we see that $y(x+1)\cdot y(x)$ is of period unity; the same holds for the function $e^{2\pi i x}$. Hence, if we multiply (16) by $y(x+1)$ the second term of its left member becomes a function of period unity. Therefore also the first term does, i.e., $y(x+1)^2\cdot y'(x)^2$ is of period unity. Hence, according to (17),

$$y(x)^2\cdot y'(x+1)^2 - y(x+1)^2\cdot y'(x)^2 = 0.$$

Thus for every solution of Σ, we have either

$$y(x)\cdot y'(x+1) - y(x+1)\cdot y'(x) = 0, \tag{18}$$

or

$$y(x)\cdot y'(x+1) + y(x+1)\cdot y'(x) = 0. \tag{19}$$

The only solution of Σ satisfying both (18) and (19) is

$$y(x) \equiv 0. \tag{20}$$

Omitting this solution, we get from (18) and (17)

$$\frac{y(x+1)}{y(x)} = \text{const.} = \pm 1.$$

The sign $+$ leads to no solution of (16); the sign $-$ gives the solutions

$$y(x) = \pm\, i\cdot e^{\pi i x} \tag{21}$$

of the system (16), (17), (18).

From (19), on the other hand, we obtain $y(x+1)\cdot y(x)$ as constant and, (20) being omitted, we may write

$$y(x+1)\cdot y(x) = \frac{1}{c^2}\cdot \tag{22}$$

Then (16), multiplied by $y(x)$, gives the solution

$$y(x) = c_1 \cdot e^{\pm c e^{\pi i x}} \qquad (c_1 = \text{const.})$$

and from (22) we finally get $c_1^2 = 1/c^2$, hence

$$y(x) = \frac{1}{c} \cdot e^{\pm c e^{\pi i x}} \tag{23}$$

with constant $c \neq 0$, as the solutions of the system (16), (17), (19).

(20), (21) and (23) are thus obtained as the manifold of Σ.

Determining $y'(x+1) \cdot y'(x)$ for each solution, we can decompose Σ into the set of the following three systems, none of which is held by any other.

$$(24) \qquad \begin{aligned} \Sigma_1 &: \quad (16), (17) \text{ and } y'(x+1) \cdot y'(x) = 0. \\ \Sigma_2 &: \quad (16), (17) \text{ and } y'(x+1) \cdot y'(x) + \pi^2 e^{2\pi i x} = 0. \\ \Sigma_3 &: \quad (16), (17) \text{ and } y'(x+1) \cdot y'(x) - \pi^2 e^{2\pi i x} = 0. \end{aligned}$$

The manifolds of Σ_1, Σ_2, Σ_3 are (20), (21), (23) respectively.

We shall show that these three systems are irreducible, so that (24) is a decomposition of Σ into essential irreducible systems.

We notice that all functions in our field $\mathfrak{F}$ are of period unity. Hence if a form F in y, with coefficients in $\mathfrak{F}$, has the solution $y(x)$, then also $y(x+1)$ is a solution of F.

Now, the irreducibility of Σ_1 is obvious. As for Σ_2, if a form F vanishes for one solution in (21), say $y(x) = +ie^{\pi i x}$, it vanishes also for $y(x+1) = -ie^{\pi i x}$, which is the other solution in (21). This shows the irreducibility of Σ_2.

Similarly, let A and B be forms such that $A \cdot B$ holds Σ_3. Then by putting

$$y(x) = \frac{1}{c} e^{+c e^{\pi i x}}$$

in $A \cdot B$, this product vanishes identically in x and c so that either A or B does. If now A vanishes for all solutions

$$y(x) = \frac{1}{c} e^{+c e^{\pi i x}}$$

then it must also vanish for

$$y(x+1) = \frac{1}{c} e^{-c e^{\pi i x}}$$

which are the other solutions in (23). Thus A holds Σ_3 and the irreducibility of Σ_3 is proved.

2. The following example shows that a pure differential equation, irreducible within the domain of differential equations (see Introduction), can be reducible, considered as a mixed difference equation.

Let $\mathfrak{F}$ be a field of constants, to be determined later, but containing π^2. The form

$$A = y'^2 + \frac{\pi^2}{4} y^2$$

is algebraically irreducible if $i\pi$ does not belong to $\mathfrak{F}$. It can be shown, however, that in this case A is also irreducible as a differential equation.

Now the solutions of A are

$$y(x) = c_1 \cdot e^{\pi i x/2} \tag{25}$$

and

$$y(x) = c_2 \cdot e^{-\pi i x/2}, \tag{26}$$

with c_1 and c_2 arbitrary constants.

For the solutions (25) we have

$$y(x+1) - i \cdot y(x) = 0, \tag{27}$$

and for the solutions (26),

$$y(x+1) + i \cdot y(x) = 0. \tag{28}$$

The only solution which (27) and (28) have in common is $y(x) \equiv 0$. Hence if i belongs to $\mathfrak{F}$ the system A is reducible, considered as a mixed difference equation.

In order to reach our purpose, we have to choose the field of numbers $\mathfrak{F}$ in such a way that it contains i and π^2 but not πi. Such a field can be obtained, for instance, by adjoining i and π^2 to the field of rational numbers.

Columbia University,
New York, N.Y.

LINEAR TRANSFORMATIONS BETWEEN HILBERT SPACES AND THE APPLICATION OF THIS THEORY TO LINEAR PARTIAL DIFFERENTIAL EQUATIONS*

BY

F. J. MURRAY

INTRODUCTION

Let us consider the equation

$$\text{(A)} \quad \begin{aligned} L(u) \equiv A_{2,0}\frac{\partial^2 u}{\partial x^2} + A_{1,1}\left(\frac{\partial^2 u}{\partial x \partial y} + \frac{\partial^2 u}{\partial y \partial x}\right) + A_{0,2}\frac{\partial^2 u}{\partial y^2} + A_{1,0}\frac{\partial u}{\partial x} \\ + A_{0,1}\frac{\partial u}{\partial y} + A_{0,0}u = v, \end{aligned}$$

where the A's and u are complex-valued functions, such that the above expression has a meaning on a bounded connected region S in the real XY-plane. We assume that the A's are bounded measurable functions on S. We shall restrict u in such a manner that (A) may be considered a linear transformation between two Hilbert spaces.

In the first five sections of this paper, we study the linear transformations between two abstract Hilbert spaces $\mathfrak{H}_1$ and $\mathfrak{H}_2$, which are regarded as coinciding only in special cases. While our investigations are naturally based on the modern work† on the subject, they are particularly closely allied to a paper of J. von Neumann.‡ Roman numerals indicate essentially new results.

In §1 and §2, we adapt the treatment of (N) to our problem to obtain the elementary theory of such transformations. In §2, we also discuss the significance in terms of groups of the adjoint. §3 deals with continuous linear transformations and while most of these results are well known,|| Theorem I is new and is used later. It also can be considered as indicating the "graphical interpretation" of certain general results of J. von Neumann, which are cited in connection with the theorem. Theorem II deals with the solution of

* Presented to the Society, March 31, 1934; received by the editors April 17 and May 3, 1934.

† Cf. Stone, *Linear Transformations in Hilbert Space*, American Mathematical Society Colloquium Publications, vol. XV, New York, 1932, also Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932. These monographs will be denoted by (S) and (B) respectively.

‡ Annals of Mathematics, (2), vol. 33 (1932), pp. 294–310. We will refer to this Memoir by (N).

|| Cf. (B), Chap. I, p. 23, Chap. III, p. 37, and Chap. VI, p. 100.

the equation $Tf=g$ and the note on constructions indicates the computational methods to be used.

We define "breakage" in §5. The theory of the manifolds which break a closed linear transformation, with domain everywhere dense in $\mathfrak{H}_1$, is shown to be completely analogous to the theory of the manifolds which reduce a self-adjoint transformation. Considerations of breakage lead to methods which, in a certain sense, compensate in the study of inverses of contractions of a closed linear transformation for the lack of compactness of Hilbert space.

In the remaining sections, we return to the consideration of (A). In §6, we set up a Hilbert space $\mathfrak{W}$, in such a manner that (A) can be considered as a limited transformation from $\mathfrak{W}$ to $\mathfrak{L}_2$ (§7). In §8, we apply Theorems 1.16 and II, to obtain a generalization of the well known methods of Ritz.* In the remaining sections we show that it is possible to use the same notions, concerning transformations between Hilbert spaces, to solve a quite general type of boundary problem (cf. §10). Certain restrictions on the boundary of S are used and these are given in §9.

Our methods apply to partial differential equations as such and we are not restricted to differential systems as previous writers have been, who use Hilbert space.† In this connection, it may be pointed out that while in the methods used here it is not necessary to take into account the distinction between elliptic, hyperbolic, and parabolic equations, nevertheless when these methods are applied, the results are essentially different. For instance, examples indicate that if the expression $L(u)$ is elliptic, the range of the transformation T of §7 is the whole space $\mathfrak{L}_2$ and that this is not true in general if $L(u)$ is non-elliptic, although the range of T may be everywhere dense in $\mathfrak{L}_2$. See examples in §8.

1. General definitions and theorems

The present section is devoted to general definitions and theorems. A knowledge of the definitions and elementary theorems of Hilbert space is assumed. (Cf. e.g. (S), Chap. 1.) We shall employ Stone's notation systematically. We shall let $(\ ,\ )_1$, $(\ ,\ )_2$ stand for the inner products in $\mathfrak{H}_1$ and $\mathfrak{H}_2$ respectively.

Definition 1.1. *By the (cartesian) product, $\mathfrak{H}_1\times\mathfrak{H}_2$, of the Hilbert spaces $\mathfrak{H}_1$ and $\mathfrak{H}_2$ (distinct or not), is meant the space of all ordered pairs $\{f_1, f_2\}$, $f_1\epsilon\mathfrak{H}_1$, $f_2\epsilon\mathfrak{H}_2$, with the linear operations $(+, a\cdot)$ and the inner product $(\ ,\ )_{1,2}$ defined by the equations*

* Journal für Mathematik, vol. 135 (1908–09), pp. 1–61.

† For an application of the methods of this space to differential equations, compare (S) (see index).

$$\{f_1, f_2\} + \{g_1, g_2\} = \{f_1 + g_1, f_2 + g_2\};$$
$$a\cdot\{f_1, f_2\} = \{af_1, af_2\};$$
$$(\{f_1, f_2\}, \{g_1, g_2\})_{1,2} = (f_1, g_1)_1 + (f_2, g_2)_2.$$

It is known that the space $\mathfrak{H}_1\times\mathfrak{H}_2$ is a Hilbert space (cf. (S), Theorem 1.26). To every pair $\{f_1, f_2\}\epsilon\mathfrak{H}_1\times\mathfrak{H}_2$ we make correspond the pair $\{f_2, f_1\}\epsilon\mathfrak{H}_2\times\mathfrak{H}_1$ and if $S\subseteq\mathfrak{H}_1\times\mathfrak{H}_2$ is given we denote by S^{-1} the set which is in $\mathfrak{H}_2\times\mathfrak{H}_1$ and whose elements correspond to elements in S.

We now introduce the notion of a many-one correspondence between a set $\mathfrak{D}\subseteq\mathfrak{H}_1$ and a set $\mathfrak{R}\subseteq\mathfrak{H}_2$ in abstract form.

DEFINITION 1.2. *Let $\mathfrak{T}$ be a non-empty subset of $\mathfrak{H}_1\times\mathfrak{H}_2$, no two of the ordered pairs of $\mathfrak{T}$ having the same first element. Then $\mathfrak{T}$ is called a transformation from $\mathfrak{H}_1$ to $\mathfrak{H}_2$; when regarded as carrying the first element f_1 of each pair $\{f_1, f_2\}\epsilon\mathfrak{T}$ into the second f_2, it is denoted by $T:f_2=Tf_1$ (T being obviously "single-valued"). By the domain $\mathfrak{D}$ of $\mathfrak{T}$ (or T) is meant the class of first elements, and by the range $\mathfrak{R}$ of $\mathfrak{T}$, the class of second elements of the class of ordered pairs $\mathfrak{T}$. We also write $\mathfrak{R}=T\mathfrak{D}$.*

When several transformations (T_1, T_2, T', T^*, etc.) are studied, their respective domains, ranges, etc., are designated with corresponding marks ($\mathfrak{D}_1$, $\mathfrak{D}_2$, $\mathfrak{D}'$, $\mathfrak{D}^*$, etc., $\mathfrak{R}_1$, $\mathfrak{R}_2$, $\mathfrak{R}'$, $\mathfrak{R}^*$, etc., $\mathfrak{T}_1$, $\mathfrak{T}_2$, $\mathfrak{T}'$, $\mathfrak{T}^*$, etc.).

DEFINITION 1.3. *Let T_1 and T_2 be two transformations from $\mathfrak{H}_1$ to $\mathfrak{H}_2$;* (a) *if $\mathfrak{T}_1=\mathfrak{T}_2$, we write $T_1=T_2$, then $\mathfrak{D}_1=\mathfrak{D}_2$ and $\mathfrak{R}_1=\mathfrak{R}_2$;* (b) *if $\mathfrak{D}_1\cdot\mathfrak{D}_2\neq 0$, T_1+T_2 shall be the transformation (cf. Definition* 1.2) *corresponding to the set $\mathfrak{T}\subseteq\mathfrak{H}_1\times\mathfrak{H}_2$, defined as follows: $\mathfrak{T}$ is the class of all pairs $\{f_1, f_2\}$ such that $f_1\epsilon\mathfrak{D}_1\cdot\mathfrak{D}_2$ and $f_2=\mathfrak{T}_1f_1+T_2f_1$;* (c) *if a is a scalar, aT_1 shall correspond to the class $\mathfrak{T}$ of all pairs $\{f_1, af_2\}\epsilon\mathfrak{H}_1\times\mathfrak{H}_2$, such that $\{f_1, f_2\}\epsilon\mathfrak{T}_1$. By Definition* 1.2, *$aT_1$ is a transformation.*

DEFINITION 1.4. *Let $\mathfrak{H}_1$, $\mathfrak{H}_2$, $\mathfrak{H}_3$ be three Hilbert spaces (distinct or not) and let T_1 and T_2 denote transformations from $\mathfrak{H}_1$ to $\mathfrak{H}_2$ and from $\mathfrak{H}_2$ to $\mathfrak{H}_3$, respectively, corresponding with the respective sets $\mathfrak{T}_1\subseteq\mathfrak{H}_1\times\mathfrak{H}_2$ and $\mathfrak{T}_2\subseteq\mathfrak{H}_2\times\mathfrak{H}_3$. Suppose furthermore $\mathfrak{R}_1\cdot\mathfrak{D}_2\neq 0$. Then the product $T=T_2T_1$ shall denote the transformation from $\mathfrak{H}_1$ to $\mathfrak{H}_3$, determined by the set $\mathfrak{T}\subseteq\mathfrak{H}_1\times\mathfrak{H}_3$, defined in the following manner. $\mathfrak{T}$ shall be the class of all pairs $\{f_1, f_3\}\epsilon\mathfrak{H}_1\times\mathfrak{H}_3$ for which there exists an element $f_2\epsilon\mathfrak{R}_1\cdot\mathfrak{D}_2(\subseteq\mathfrak{H}_2)$ such that $\{f_1, f_2\}\epsilon\mathfrak{T}_1$ and $\{f_2, f_3\}\epsilon\mathfrak{T}_2$.*

It should be observed that the symbol T_1T_2 has no meaning in general, and even in the special case where $\mathfrak{H}_1=\mathfrak{H}_2=\mathfrak{H}_3$ and T_1T_2 can exist, it is, in general, distinct from T_2T_1.

DEFINITION 1.5. *The notation* $T_1 \supseteq T_2$ *and* $T_1 \supset T_2$ *is used to signify that* $\mathfrak{T}_1 \supseteq \mathfrak{T}_2$ *and* $\mathfrak{T}_1 \supset \mathfrak{T}_2$, *respectively. In the former case* T_1 *is said to be an extension of* T_2, T_2 *a contraction of* T_1, *or the contraction of* T_1 *with domain* $\mathfrak{D}_2$; *in the latter case* T_1 *is said to be a proper extension of* T_2, T_2 *a proper contraction of* T_1.

DEFINITION 1.6. *Let the transformation* T *from* $\mathfrak{H}_1$ *to* $\mathfrak{H}_2$ *be such that the set* $\mathfrak{T}$ *contains no two pairs with the same last elements. Then* T *is said to have an inverse* T^{-1}, *which is defined by the set* $\mathfrak{T}^{-1} \subseteq \mathfrak{H}_2 \times \mathfrak{H}_1$, *consisting of the pairs in* $\mathfrak{T}$, *with their order inverted.*

Note the obvious relations

$$\mathfrak{D}^{-1} = \mathfrak{R};\ \mathfrak{R}^{-1} = \mathfrak{D};\ TT^{-1} \subseteq I,\ \ T^{-1}T \subseteq I.$$

DEFINITION 1.7. *A sequence of transformations* $\{T_n\}$ *from* $\mathfrak{H}_1$ *to* $\mathfrak{H}_2$, *with domains* $\{\mathfrak{D}_n\}$, *is said to converge on the set* $\mathfrak{S}$ *if*

(1) $\mathfrak{S} \subseteq \liminf \mathfrak{D}_n$;*

(2) *for every* $f \in \mathfrak{S}$, *the sequence* $\{f, T_n f\}$ *converges.*

A sequence of transformations $\{T_n\}$, *from* $\mathfrak{H}_1$ *to* $\mathfrak{H}_2$, *is said to have the limit* T *on* $\mathfrak{S}$ *if*

(1) $\mathfrak{S} \subseteq (\liminf \mathfrak{D}_n) \cdot \mathfrak{D}$;

(2) *for every* $f \in \mathfrak{S}$, *the sequence* $\{f, T_n f\}$ *converges to* $\{f, Tf\}$.

We shall write $T_n \rightarrow T$, $n \rightarrow \infty$ on $\mathfrak{S}$.

Since Hilbert space is complete, it is easy to show

THEOREM 1.1. *If a sequence* $\{T_n\}$ *from* $\mathfrak{H}_1$ *to* $\mathfrak{H}_2$ *converges on* $\mathfrak{S}$, *then there exists a transformation* T *from* $\mathfrak{H}_1$ *to* $\mathfrak{H}_2$, *such that* $T_n \rightarrow T$, $n \rightarrow \infty$ *on* $\mathfrak{S}$. *If* $T_n \rightarrow T$ *on* $\mathfrak{S}$, *then the sequence* $\{T_n\}$ *is convergent on* $\mathfrak{S}$.

DEFINITION 1.8. *A transformation* T *from* $\mathfrak{H}_1$ *to* $\mathfrak{H}_2$ *is said to be continuous at an element* f *in its domain, if to each positive number* ϵ, *there corresponds a positive number* $\delta = \delta(\epsilon, f)$ *such that whenever* g *is an element of* $\mathfrak{D}$, *satisfying the inequality* $\|f-g\|_1 \leqq \delta$, *the element* Tg *satisfies the inequality* $\|Tf-Tg\|_2 \leqq \epsilon$. *A transformation* T *is said to be continuous, if it is continuous at every element in its domain.*

DEFINITION 1.9. *A transformation* T *from* $\mathfrak{H}_1$ *to* $\mathfrak{H}_2$ *is said to be closed if* $\mathfrak{T}$ *is closed.*

DEFINITION 1.10. *A transformation* T *from* $\mathfrak{H}_1$ *to* $\mathfrak{H}_2$ *is said to be linear if* $\mathfrak{T}$ *is a linear manifold.*

---

* $\liminf \mathfrak{D}_n$ is the set of points, each of which is in all but a finite number of the ("almost every") $\mathfrak{D}_n$.

We state the following theorems the proofs of which are easy and may be omitted.

THEOREM 1.2. *The domain and range of a linear transformation are linear manifolds.*

THEOREM 1.3. *If T_1 and T_2 are two linear transformations from $\mathfrak{H}_1$ to $\mathfrak{H}_2$, then aT_1 and T_1+T_2 are linear transformations.*

THEOREM 1.4. *If T_2 is a linear transformation from $\mathfrak{H}_1$ to $\mathfrak{H}_2$ and T_1 is a linear transformation from $\mathfrak{H}_2$ to $\mathfrak{H}_3$, then $T=T_1T_2$ is linear.*

THEOREM 1.5. *If T is a linear transformation possessing an inverse then T^{-1} is linear. If T is a closed transformation possessing an inverse, then T^{-1} is closed.*

THEOREM 1.6. *If T is a transformation, any subset of $\mathfrak{T}$ is a transformation.*

THEOREM 1.7. *If the transformation T has a linear extension then there exists a unique linear transformation $\hat{T}$ such that*

(1) *$\hat{T}$ is an extension of T;*

(2) *if T' is linear and $T' \supseteq T$, then $T' \supseteq \hat{T}$.*

If T has a closed linear extension then there exists a unique closed linear transformation $\tilde{T}$ such that

(1) *$\tilde{T}$ is an extension of T and $\hat{T}$;*

(2) *if T' is a closed linear extension of T, $T' \subseteq \tilde{T}$.*

If T has a (closed) linear extension T_0, we take for $\hat{T}$ (or $\tilde{T}$) the (closed) linear manifold determined by $\mathfrak{T}$. The theorem follows immediately from Theorem 1.6, Definitions 1.8 and 1.9 above, and (S) Definition 1.4.

2. THE PERPENDICULAR AND ADJOINT OF A TRANSFORMATION

We return to the study of perfectly general transformations T_1 from $\mathfrak{H}_1$ to $\mathfrak{H}_2$ and T_2 from $\mathfrak{H}_2$ to $\mathfrak{H}_1$, determined by the sets $\mathfrak{T}_1$ and $\mathfrak{T}_2$ in $\mathfrak{H}_1 \times \mathfrak{H}_2$ and $\mathfrak{H}_2 \times \mathfrak{H}_1$ respectively While our treatment paraphrases the work of (N), we wish to call attention to the simplicity in conceptions and proofs obtained by studying directly not the adjoint but the perpendicular to a transformation, which notion we define as follows.

DEFINITION 1.11. *The transformations T_1 from $\mathfrak{H}_1$ to $\mathfrak{H}_2$, and T_2 from $\mathfrak{H}_2$ to $\mathfrak{H}_1$, shall be said to be perpendicular (symbolically $T_1 \perp T_2$) if $\mathfrak{T}_1^{-1}$ and $\mathfrak{T}_2$ are orthogonal in $\mathfrak{H}_2 \times \mathfrak{H}_1$. They shall be said to be adjoint ($T_1 \wedge T_2$), provided that T_1 and $-T_2$ are perpendicular.*

Obviously $T_1 \perp T_2$ implies $T_2 \perp T_1$; also $T_1 \perp T_2$ implies that, for all pairs $\{f, T_1f\} \epsilon \mathfrak{T}$ and $\{T_2g, g\} \epsilon \mathfrak{T}_2$,

$$0 = (\{f, T_1f\}, \{T_2g, g\})_{1,2} = (f, T_2g)_1 + (T_1f, g)_2,$$

and conversely this equation, valid for all $f\epsilon\mathfrak{D}_1$ and $g\epsilon\mathfrak{D}_2$, implies that $T_1 \perp T_2$. We are immediately led to the following theorem which connects our definition with the usual one.†

THEOREM 1.8. *The relation $T_1 \wedge T_2$ will hold if and only if*

$$(f, T_2g)_1 = (T_1f, g)_2$$

for all $f\epsilon\mathfrak{D}_1$, and $g\epsilon\mathfrak{D}_2$, and further $T_1 \wedge T_2$ implies $T_2 \wedge T_1$.

We now suppose given a perfectly general transformation T from $\mathfrak{H}_1$ to $\mathfrak{H}_2$, $\mathfrak{T}$ and $\mathfrak{D}$ being as usual, and consider the problem of finding transformations perpendicular to it. We are led to the following theorem.

THEOREM 1.9. *Let $\mathfrak{T}^\perp$ be the class of all elements of $\mathfrak{H}_2 \times \mathfrak{H}_1$, orthogonal to $\mathfrak{T}^{-1}$. Then a necessary and sufficient condition that $\mathfrak{T}^\perp$ constitute a transformation from $\mathfrak{H}_2$ to $\mathfrak{H}_1$, is that $\mathfrak{D}$ span $\mathfrak{H}_1$.‡ When $\mathfrak{D}$ spans $\mathfrak{H}_1$, then $T \perp T'$ if and only if $\mathfrak{T}' \subseteq \mathfrak{T}^\perp$.*

The proof is quite analogous to that of (S) Theorem 2.6.

DEFINITION 1.12. *When the transformation T from $\mathfrak{H}_1$ to $\mathfrak{H}_2$ is such that $\mathfrak{D}$ spans $\mathfrak{H}_1$, then $T^\perp$ defined by the set $\mathfrak{T}^\perp$, of Theorem 1.9, is called the perpendicular of T, and $T^* = -T^\perp$ is called the adjoint of T.*§

The following Theorems are capable of very simple proofs which we omit.

THEOREM 1.10. *If $T^\perp$ exists, then $\mathfrak{T}^\perp$ is the orthogonal complement in $\mathfrak{H}_2 \times \mathfrak{H}_1$ of the closed linear manifold determined by $\mathfrak{T}^{-1}$.*

THEOREM 1.11. *If $T^\perp$ exists, then it is a closed linear transformation.*

THEOREM 1.12. *If T is a transformation from $\mathfrak{H}_1$ to $\mathfrak{H}_2$, whose domain determines $\mathfrak{H}_1$, and whose range determines $\mathfrak{H}_2$, and if T possesses an inverse, then $T^\perp$ and $(T^{-1})^\perp$ are inverses.*

$\mathfrak{T}^\perp$ and $(\mathfrak{T}^{-1})^\perp$ correspond to transformations by Theorem 1.9 and it is easy to verify that $(\mathfrak{T}^\perp)^{-1} = (\mathfrak{T}^{-1})^\perp$, since the correspondence $\mathfrak{A} \sim \mathfrak{A}^{-1}$ preserves orthogonality.

† Cf. (S) Definition 2.7, (B) p. 99.

‡ The closed linear manifold determined by $\mathfrak{D}$ is $\mathfrak{H}_1$.

§ For limited transformations (cf. §4), in view of Theorem 1.8, this definition is equivalent to that given by (B) p. 99. Remembering that a Hilbert space is self-dual, we see that to every linear functional on $\mathfrak{H}_2$, Y or on $\mathfrak{H}_1, X$, we can make correspond respectively an $f\epsilon\mathfrak{H}_2$ and an $f^*\epsilon\mathfrak{H}_1$ such that $Y(h) = (h, f)_2$; $X(g) = (g, f^*)_1$ for all h in $\mathfrak{H}_2$ and all g in $\mathfrak{H}_1$. Thus when T is limited and $X = \overline{T}Y$ then $(g, f^*)_1 = X(g) = \overline{T}(Y)(g) = Y(Tg) = (Tg, f)_2$ or $f^* = T^*f$.

THEOREM 1.13. *A necessary and sufficient condition that a transformation T from $\mathfrak{H}_1$ to $\mathfrak{H}_2$, with domain everywhere dense in $\mathfrak{H}_1$, should have a closed linear extension is that the domain of $T^\perp$ (and hence of T^*) be everywhere dense in $\mathfrak{H}_2$. If $\tilde{T}$ exists then $\tilde{T}=(T^\perp)^\perp=-(-T^*)^*=(T^*)^*$.*

The proofs of this theorem and the three following are quite similar to the proofs of the corresponding theorems in the memoir of von Neumann.

THEOREM 1.14. *Let T be a closed linear transformation from $\mathfrak{H}_1$ to $\mathfrak{H}_2$, with domain everywhere dense in $\mathfrak{H}_1$. Then $T^\perp$ and T^* are closed and linear and their common domain is everywhere dense in $\mathfrak{H}_2$, and T^*T is self-adjoint.*†

THEOREM 1.15. *Let T be as in Theorem 1.14. Let $\mathfrak{D}_1$ be the domain of T^*T and T_1 the contraction of T with domain $\mathfrak{D}_1$. Then $\tilde{T}_1=T$.*

THEOREM 1.16. *Let T be as in Theorem 1.14. Let the linear manifold of all $f\epsilon\mathfrak{H}_1$ such that $Tf=0$ be denoted by $\mathfrak{N}$ and the linear manifold of all $f\epsilon\mathfrak{H}_2$ such that $T^*f=0$ be denoted by $\mathfrak{N}^*$; then $\mathfrak{N}$ and $\mathfrak{N}^*$ are closed, and if $\overline{\mathfrak{M}}(\mathfrak{N})$ and $\overline{\mathfrak{M}}(\mathfrak{N}^*)$ denote the closed linear manifolds determined by $\mathfrak{N}$ and $\mathfrak{N}^*$, respectively, then*

$$\overline{\mathfrak{M}}(\mathfrak{N}^*) = \mathfrak{H}_1 \ominus \mathfrak{N}; \quad \overline{\mathfrak{M}}(\mathfrak{N}) = \mathfrak{H}_2 \ominus \mathfrak{N}^*.$$

A Hilbert space is of course an Abelian group with operators (i.e., multiplication by scalars).‡ A linear manifold $\mathfrak{M}$ is an Abelian subgroup of $\mathfrak{H}$ and hence we can form the quotient group $\mathfrak{H}/\mathfrak{M}$, i.e., the group of the remainder classes of $\mathfrak{H}$, mod $\mathfrak{M}$.§ Now if $\mathfrak{M}$ is closed, $\mathfrak{H}\ominus\mathfrak{M}$ is simply isomorphic to $\mathfrak{H}/\mathfrak{M}$ under the correspondence which links $f\epsilon\mathfrak{H}\ominus\mathfrak{M}$ with the remainder class, which consists of all elements congruent to f, mod $\mathfrak{M}$.

Let $\mathfrak{T}$ correspond to a closed linear transformation from $\mathfrak{H}_1$ to $\mathfrak{H}_2$. $\mathfrak{T}^\perp\subseteq\mathfrak{H}_2\times\mathfrak{H}_1$ is simply isomorphic to the quotient group of $\mathfrak{T}^{-1}$. These notions persist even when considered in linear vector space $\mathfrak{E}$ more general than Hilbert space where, however, $\mathfrak{T}^\perp$ is isomorphic with the set of linear functionals on $\mathfrak{E}_1\times\mathfrak{E}_2/\mathfrak{T}^{-1}$.

3. CONTINUOUS LINEAR TRANSFORMATIONS

We now consider continuous linear transformations. Where proofs are omitted the theorems are either special cases of theorems in (B) or their proofs are quite analogous to the proofs of corresponding theorems in (S).

THEOREM 1.17. *If a linear transformation T from $\mathfrak{H}_1$ to $\mathfrak{H}_2$ is continuous at one element in its domain it is uniformly continuous at each point in its domain.*

† Cf. (S) 2.11. A transformation (in $\mathfrak{H}$) is said to be self-adjoint if it is equal to its adjoint.

‡ Cf. van der Waerden, *Höhere Algebra*, vol. 1, chap. 6, Springer, 1932.

§ Cf. van der Waerden, loc. cit., vol. 1, pp. 31–36, 132–135.

THEOREM 1.18. *If T is a linear transformation from $\mathfrak{H}_1$ to $\mathfrak{H}_2$, then T is continuous if and only if there is a positive number C such that for all $f\epsilon\mathfrak{D}$,*

$$\|Tf\|_2 \leqq C\|f\|_1.$$

DEFINITION 1.13. *A linear transformation T from $\mathfrak{H}_1$ to $\mathfrak{H}_2$ is said to be limited if there exists a C such that for all $f\epsilon\mathfrak{D}$*

$$\|Tf\|_2 \leqq C\|f\|_1.$$

The least such C is called the bound of C.

THEOREM 1.19. *If T is a continuous linear transformation from $\mathfrak{H}_1$ to $\mathfrak{H}_2$, with bound C, then $\tilde{T}$ exists and has the bound C and the domain of $\tilde{T}$ is the closed linear manifold determined by $\mathfrak{D}$.*

THEOREM 1.20. *If T is a limited transformation with bound C from $\mathfrak{H}_1$ to $\mathfrak{H}_2$, with domain spanning $\mathfrak{H}_1$, then T^* exists and is closed and limited with bound C and domain $\mathfrak{H}_2$.*

THEOREM I. *Let T be a closed limited transformation from $\mathfrak{H}_1$ to $\mathfrak{H}_2$ with bound C. Let $\{f_n\}$ be a sequence everywhere dense in $\mathfrak{D}$. The sequences $\{f_n, Tf_n\}$ and $\{Tf_n\}$ are everywhere dense in $\mathfrak{T}$ and $\mathfrak{R}$ respectively. Let $\{\phi_n\}$ be a sequence which determines $\mathfrak{D}$. Then the sequences $\{\phi_n, T\phi_n\}$ and $\{T\phi_n\}$ determine $\mathfrak{T}$ and $\mathfrak{R}$ respectively.*

Let $\{f, Tf\}\epsilon\mathfrak{T}$. Then given an $\epsilon>0$, we can find an f such that

$$\|f_n - f\| < \frac{\epsilon}{1+C}.$$

Then

$$\begin{aligned}\|\{f, Tf\} - \{f_n, Tf_n\}\|_{1,2} &= \|\{f - f_n, T(f - f_n)\}\|_{1,2} \\ &= (\|f - f_n\|_1^2 + \|Tf - Tf_n\|_2^2)^{1/2} \leqq \|f - f_n\|_1 + \|Tf - Tf_n\|_2 \\ &= \|f - f_n\|_1 + \|T(f - f_n)\|_2 \leqq (1 + C)\|f - f_n\|_1 < \epsilon.\end{aligned}$$

We have also shown that $\|Tf-Tf_n\|<\epsilon$.

The proof of the last statement of the theorem is simplified by the following lemma whose proof is almost immediate. It is convenient to introduce the following notation. We denote by $\mathfrak{M}(\mathfrak{S})$ the linear manifold determined by $\mathfrak{S}$, and by $\overline{\mathfrak{M}}(\mathfrak{S})$ the closed linear manifold determined by $\mathfrak{S}$.

LEMMA. *A necessary and sufficient condition that a sequence $\{k_n\}$ determine a closed linear manifold $\mathfrak{D}$, is that the denumerable set $R(\{k_n\})$ consisting of all elements in the form $\sum_1^n r_i k_i$, n being any integer, the r_i being rational complex, be everywhere dense in $\mathfrak{D}$.*

We apply the lemma to show that $R(\{\phi_i\})$ is everywhere dense in $\mathfrak{D}$. Then we apply the part of the theorem already proved to show that $R(\{\{\phi_i, T\phi_i\}\})$ and $R(\{T\phi_i\})$ are everywhere dense in $\mathfrak{T}$ and $\mathfrak{R}$ respectively. These facts and the lemma imply that $\{\{\phi_i, T\phi_i\}\}$ and $\{T\phi_i\}$ determine $\mathfrak{T}$ and $\mathfrak{R}$, respectively.

THEOREM II. *Let T be a limited transformation from $\mathfrak{H}_1$ to $\mathfrak{H}_2$, with domain $\mathfrak{H}_1$ and range determining $\mathfrak{H}_2$. Then T^* exists and is limited with domain $\mathfrak{H}_2$ and has an inverse. Let $\{\psi_i\}$ be a complete orthonormal set in $\mathfrak{H}_1$, $\{\phi_i\}$ a complete orthonormal set in $\mathfrak{H}_2$.*

Let $\{\chi_i\}$ be a sequence in $\mathfrak{H}_1$, which is obtained by orthonormalizing the sequence $\{T^*\phi_i\}$. Then the sequence $\{\chi_i\}$ determines $\mathfrak{H}_1 \ominus \mathfrak{N}$; or $Tf=0$ if and only if $(f, \chi_i)_1=0$ for every i. A necessary and sufficient condition that an element $g \epsilon \mathfrak{H}_2$ be in the range of T is that

$$\sum_1^\infty |(g, T^{*-1}\chi_i)_2|^2 < \infty .$$

If $g \epsilon \mathfrak{R}$ and if

$$f = \sum_1^\infty (g, T^{*-1}\chi_i)_2 \chi_i ,$$

then $Tf=g$.

That T^* exists and is limited with domain $\mathfrak{H}_2$ is implied by Theorem 1.20. Since the manifold of zeros of T^*, $\mathfrak{N}^*$, is by Theorem 1.16 equal to $\mathfrak{H}_2 \ominus \overline{\mathfrak{M}}(\mathfrak{R})$, $\mathfrak{N}^*$ contains only the element 0, since $\overline{\mathfrak{M}}(\mathfrak{R}) = \mathfrak{H}_2$, by hypothesis. Hence T^{*-1} exists.

Since the sequence $\{\phi_i\}$ determines $\mathfrak{H}_2$ and T^* is limited, it follows from Theorem I that the set $\{T^*\phi_i\}$ determines $\overline{\mathfrak{M}}(\mathfrak{R}^*)$ which is $\mathfrak{H}_1 \ominus \mathfrak{N}$, by Theorem 16. Hence when we orthonormalize the set $\{T^*\phi_i\}$, the resulting sequence determines $\mathfrak{H}_1 \ominus \mathfrak{N}$.

Now one can readily verify that $\{\chi_i\} \subseteq \mathfrak{D}^{*-1}$ and that $\{T^{*-1}\chi_i\}$ determines $\mathfrak{H}_2$ since when we orthonormalize $\{T^{*-1}\chi_i\}$ we obtain $\{\phi_i\}$.

Now if $g \epsilon \mathfrak{R}$, there exists an $f' \epsilon \mathfrak{H}_1$, such that $g = Tf'$. If E is the projection with range $\overline{\mathfrak{M}}(\mathfrak{R}^*)$, then $I-E$ is the projection with range $\mathfrak{N}$, the manifold of zeros of T, by Theorem 1.16. Hence $g = Tf' = TEf' + T(I-E)f' = TEf' = Tf$, letting $Ef' = f$. Since as shown above the sequence $\{\chi_i\}$ determines $\overline{\mathfrak{M}}(\mathfrak{R}^*)$, we have

$$f = \sum_1^\infty a_i \chi_i, \qquad \sum_1^\infty |a_i|^2 < \infty ,$$

where

$$a_i = (f, \chi_i)_1 = (f, T^*T^{*-1}\chi_i)_1 = (Tf, T^{*-1}\chi_i)_2 = (g, T^{*-1}\chi_i)_2 .$$

Thus the condition stated in the theorem is necessary.

It is also sufficient. For suppose

$$\sum_1^\infty |a_i|^2 < \infty, \qquad a_i = (g, T^{*-1}\chi_i)_2.$$

Let $f=\sum_1^\infty a_i\chi_i$; we will show that $Tf=g$. For we have

$$(T^{*-1}\chi_i, T\chi_j)_2 = (T^*T^{*-1}\chi_i, \chi_j)_1 = (\chi_i, \chi_j)_1 = \delta_{ij},$$

and hence, since T is continuous, for every i

$$\begin{aligned}(g - Tf, T^{*-1}\chi_i)_2 &= \left(g - \sum_1^\infty a_jT\chi_j;\ T^{*-1}\chi_i\right)_2 \\ &= (g, T^{*-1}\chi_i)_2 - \sum_1^\infty a_j(T\chi_j, T^{*-1}\chi_i) = (g, T^{*-1}\chi_i) - a_i \\ &= (g, T^{*-1}\chi_i) - (g, T^{*-1}\chi_i) = 0.\end{aligned}$$

But we have already shown that the sequence $\{T^{*-1}\chi_i\}$ determines $\mathfrak{H}_2$. Hence $g-Tf=0$ and $g\epsilon\mathfrak{R}$.

If T is not limited and if $\{f_n\}$ is a sequence which determines $\overline{\mathfrak{M}}(\mathfrak{D})$ and each $f_n\epsilon\mathfrak{D}$, then the sequence $\{f_n, Tf_n\}$ does not necessarily determine $\mathfrak{T}$. For instance let D be the transformation in $\mathfrak{L}_2$, for the interval $0\leqq x\leqq 2\pi$ ((S) Theorem 1.24) with domain consisting of all elements in the form $f=c+\int_0^x g(\xi)d\xi$, $g\epsilon\mathfrak{L}_2$ and such that $Df=g$. The sequence $\{e^{inx}\}$, $n=0, \pm 1, \pm 2, \cdots$, determines $\mathfrak{L}_2$, but the element $\{e^x-e^{2\pi-x}, e^x+e^{2\pi-x}\}$ is orthogonal to the sequence $\{\{e^{inx}, ine^{inx}\}\}$, $n=0, \pm 1, \pm 2, \cdots$.†

Note on constructability. It is desirable for the applications which we will make later to have a method of procedure by which one can use Theorem 1.16 and Theorem II without recourse to Zermelo's Axiom. Zermelo's Axiom is used in showing that a sub-set of a separable space is separable. For the applications mentioned it is sufficient to give a method using the operations of the postulates ((S) p. 3) by which having a set which determines $\mathfrak{M}$ we obtain a set which determines $\mathfrak{H}\ominus\mathfrak{M}$. Let $\mathfrak{M}$ be the range of the projection E. Let $\{f_i\}$ determine $\mathfrak{M}$. Let $\{\chi_i\}$ be the result of applying the Gram-Schmidt process to $\{f_i\}$. Let $\{\phi_i\}$ be a complete orthonormal set in $\mathfrak{H}$. Then we notice that

$$E\phi_i = \sum_1^\infty a_i\chi_i, \qquad a_i = (\phi_i, \chi_i),$$

† Compare this example with the results in the paper by J. von Neumann, on non-limited infinite matrices, Journal für Mathematik, vol. 161 (1929), pp. 208–236. It offers a "graphical interpretation" of the general results.

and if $f \epsilon \mathfrak{H} \ominus \mathfrak{M}$ then since E is limited

$$f = (1 - E)f = (1 - E)\sum_1^\infty b_i\phi_i = \sum_1^\infty b_i(1 - E)\phi_i, \qquad b_i = (f, \phi_i).$$

Hence the sequence $\{(1-E)\phi_i\}$ determines $\mathfrak{H} \ominus \mathfrak{M}$.

Isometric transformations

Definition 1.14. *U will be called a unitary transformation from $\mathfrak{H}_1$ to $\mathfrak{H}_2$ if its domain is $\mathfrak{H}_1$ and its range is $\mathfrak{H}_2$ and if for every f and g in its domain we have*

$$(uf, ug)_2 = (f, g)_1.$$

Definition 1.15. *A linear transformation V from $\mathfrak{H}_1$ to $\mathfrak{H}_2$ will be said to be isometric if, for every f and h in its domain, we have*

$$(vf, vh)_2 = (f, h).$$

It is hardly necessary to point out that there is at least one unitary transformation between any two Hilbert spaces, since if $\{\phi_i\}$ is a complete orthonormal set in $\mathfrak{H}_1$ and if $\{\psi_i\}$ is a complete orthonormal set in $\mathfrak{H}_2$ and if $\mathfrak{T} = \{\{\phi_i, \psi_i\}\}$, then $U = \tilde{T}$ is obviously unitary.

The proofs of the following two theorems are quite similar to the proofs of (S) Theorems 2.42 and 2.43.

Theorem 1.21. *A unitary transformation U from $\mathfrak{H}_1$ to $\mathfrak{H}_2$ is linear continuous and isometric. U^* exists and is unitary. U^{-1} exists and $U^{-1} = U^*$.*

Theorem 1.22. *An isometric transformation V from $\mathfrak{H}_1$ to $\mathfrak{H}_2$ is limited and possesses an isometric inverse. The transformation $\tilde{V}$ exists and is isometric. The domain and range of $\tilde{V}$ are the closed linear manifolds determined by the domain and range of V respectively.*

4. von Neumann's theorem

By means of the previous discussion, we are now able to obtain results corresponding to the remaining Theorems in (N) in a manner analogous to the proofs given in that memoir. These results follow.

We introduce the transformation $B = (T^*T)^{1/2}$, which is of course in $\mathfrak{H}_1$.

Theorem 1.23. *If T is a closed linear transformation from $\mathfrak{H}_1$ to $\mathfrak{H}_2$, with domain everywhere dense in $\mathfrak{H}_1$, and if $B = (T^*T)^{1/2}$, B is self-adjoint in $\mathfrak{H}_1$, with domain $\mathfrak{D}$ and $(Tf, Tg)_2 = (Bf, Bg)_1$ for all f and g of $\mathfrak{D}$.*

THEOREM 1.24. *Let T be a closed linear transformation from $\mathfrak{H}_1$ to $\mathfrak{H}_2$, with domain everywhere dense in $\mathfrak{H}_1$. Let $B=(T^*T)^{1/2}$, $C=(TT^*)^{1/2}$; B and C are self-adjoint. The domain of B is $\mathfrak{D}$, the domain of C is $\mathfrak{D}^*$. Let $\mathfrak{R}_b$ denote the range of B, $\mathfrak{R}_c$ that of C. Let $\mathfrak{N}$ and $\mathfrak{N}^*$ be as in Theorem* 1.16. *$Bf=0$ if and only if $f\epsilon\mathfrak{N}$, $Cf=0$ if and only if $f\epsilon\mathfrak{N}^*$. Also*

$$\overline{\mathfrak{M}}(\mathfrak{R}_b) = \overline{\mathfrak{M}}(\mathfrak{R}^*) = \mathfrak{H}_1 \ominus \mathfrak{N};\ \overline{\mathfrak{M}}(\mathfrak{R}) = \overline{\mathfrak{M}}(\mathfrak{R}_c) = \mathfrak{H}_2 \ominus \mathfrak{N}^*.$$

Let E be the projection in $\mathfrak{H}_1$ on $\mathfrak{H}_1\ominus\mathfrak{N}=\overline{\mathfrak{M}}(\mathfrak{R}^*)$, E' be the projection in $\mathfrak{H}_2$ on $\overline{\mathfrak{M}}(\mathfrak{R})$.

There is a transformation W from $\mathfrak{H}_1$ to $\mathfrak{H}_2$, with domain $\mathfrak{H}_1$ and continuous, which takes $\overline{\mathfrak{M}}(\mathfrak{R}_b)$ in a single-valued isometric manner onto $\overline{\mathfrak{M}}(\mathfrak{R})$; W^* takes $\overline{\mathfrak{M}}(\mathfrak{R}_c)$ in a single-valued isometric manner onto $\overline{\mathfrak{M}}(\mathfrak{R}^*)$. W is zero on $\mathfrak{N}$, W^* is zero on $\mathfrak{N}^*$.

Furthermore $W^*W=E$, $WW^*=E'$,

$$T = WB = CW; \qquad T^* = BW^* = W^*C;$$
$$B = W^*T = T^*W; \qquad C = TW^* = WT^*;$$
$$C = WBW^*; \qquad B = W^*CW.$$

The following theorem is now a simple consequence of (S) Theorem 2.25.

THEOREM 1.25. *If T is a closed transformation from $\mathfrak{H}_1$ to $\mathfrak{H}_2$ with domain $\mathfrak{H}_1$, then T is bounded.*†

5. BREAKAGE

A closed linear transformation is a correspondence which is isomorphic with respect to Postulate A, (S) p. 3. We will study in this section to what extent the relationship of orthogonality, introduced in Postulate B, is preserved.

DEFINITION I. *Let T be a closed linear transformation from $\mathfrak{H}_1$ to $\mathfrak{H}_2$, with domain everywhere dense in $\mathfrak{H}_1$. A projection F in $\mathfrak{H}_1$ is said to break T if*

(1) *$Ff\epsilon\mathfrak{D}$ for all $f\epsilon\mathfrak{D}$, and*

(2) *there exists a projection F' (independent of f) such that $TFf=F'Tf$ for all $f\epsilon\mathfrak{D}$.*

If F breaks T and $\mathfrak{M}$ is the range of F, then $\mathfrak{M}$ will be said to break T.

We note that conditions (1) *and* (2) *are equivalent to $TF\supseteq F'T$.*

THEOREM III. *If F breaks T, a closed linear transformation from $\mathfrak{H}_1$ to $\mathfrak{H}_2$ with domain everywhere dense in $\mathfrak{H}_1$, then F' breaks T^* and $T^*F'f=FT^*f$ for all $f\epsilon\mathfrak{D}^*$.*

† (B) chap. III, Theorem 7, p. 41.

For all $f\epsilon\mathfrak{D}$ and $g\epsilon\mathfrak{D}^*$, we have

$$(Tf, F'g)_2 = (F'Tf, g)_2 = (TFf, g)_2 = (Ff, T^*g)_1 = (f, FT^*g).$$

But this equation implies $T^*F'g = FT^*g$ for all $g\epsilon\mathfrak{D}^*$, or $T^*F' \supseteq FT^*$.

Henceforth all quantities will have their significance as in Theorem 1.24.

THEOREM IV. *F breaks T, if and only if it reduces B. If F breaks T, then* $F' = WFW^* + F_0'$, *where* F_0' *is a projection whose range is included in* $\mathfrak{N}^* = \mathfrak{H}_2 \ominus \overline{\mathfrak{M}}(\mathfrak{R})$ *and for any such* F', $TF \supseteq F'T$.

We first show that if F reduces B, then WFW^* is a projection. WFW^* is obviously a limited symmetric transformation. We must show that $(WFW^*)^2 = WFW^*$. To do this we prove that $EFW^* = FW^*$. Let $W^{*\prime}$ be the contraction of W^* with domain $\mathfrak{R}\cdot+\cdot\mathfrak{N}^*$.† Since the domain of $W^{*\prime}$ and also W^* are linear, $W^{*\prime}$ is linear and also limited since W^* is limited.

Since, by hypothesis, F reduces B, and by Theorem 1.24, $W^*T = B = T^*W$,

$$FW^{*\prime}T = FW^*T = FB \subseteq BF = T^*WF.$$

Thus, since E is the projection on $\overline{\mathfrak{M}}(\mathfrak{R})$, we have

$$EFW^{*\prime}T = FW^{*\prime}T,$$

or $EFW^{*\prime}f = FW^{*\prime}f$, $f\epsilon\mathfrak{R}$. But for $g\epsilon\mathfrak{N}^*$, $W^{*\prime}g = W^*g = 0$ and $EFW^{*\prime}g = FW^{*\prime}g = 0$. Now an element of $\mathfrak{R}\cdot+\cdot\mathfrak{N}^*$ is the sum of an element of $\mathfrak{R}$ and an element of $\mathfrak{N}^*$, and we can conclude

$$EFW^{*\prime} = FW^{*\prime}.$$

But $\mathfrak{R}\cdot+\cdot\mathfrak{N}^*$, the domain of this transformation, is everywhere dense in $\mathfrak{H}_2$. Hence by Theorem 1.19, it has a unique closed extension and thus $EFW^* = FW^*$, since both are closed and $EFW^* \supseteq EFW^{*\prime}$ and $FW^* \supseteq FW^{*\prime}$.

Now

$$\begin{aligned}(WFW^*)^2 &= (WFW^*)(WFW^*) = (WF)(W^*W)(FW^*) = (WF)(E)(FW^*)\\ &= (WF)(EFW^*) = (WF)(FW^*) = (W)(FF)(W^*) = WFW^*.\end{aligned}$$

Hence by (S) Theorem 2.36, WFW^* is a projection. Now since the domain of B is the same as that of T, $Ff\epsilon\mathfrak{D}$ if $f\epsilon\mathfrak{D}$. We also have, since $B = W^*T$ and $E' = WW^*$ is the projection on the closed linear manifold determined by $\mathfrak{R}$,

$$TF = E'TF = (W)(W^*TF) = WBF \supseteq WFB = WFW^*T.$$

Now conversely let us suppose that F breaks T and let F' be as in Definition I. Then by Theorem III, we have

† $\mathfrak{R}\cdot+\cdot\mathfrak{N}^*$ contains every element which is the sum of an element of $\mathfrak{R}$ and an element of $\mathfrak{N}^*$ and only such elements.

$$FT^*T \subseteq T^*F'T \subseteq T^*TF.$$

One infers from this that F reduces $T^*T=BB$ and hence, by a simple application of (S) Theorem 8.1, that F reduces B.

Let F break T. Now suppose F' is such that $TF \supseteq F'T$. Then F' is in the form stated in the theorem. For we have for all $f \epsilon \mathfrak{D}$,

$$TFf = WFW^*Tf = F'Tf.$$

Hence $WFW^*f=F'f$, $f\epsilon\mathfrak{R}$.

Now let H be the contraction of WFW^*, with domain $\mathfrak{R}$. Since $\mathfrak{R}$ and WFW^* are linear, H is linear. Let K be the contraction of F', with domain $\mathfrak{R}$. It is also linear, and from the above we have that $H=K$. Applying Theorem 1.19, we have

$$\tilde{H} = \tilde{K},$$

and this transformation has domain $\overline{\mathfrak{M}}(\mathfrak{R})$. It is also a contraction of WFW^* and F'. Hence since the range of E' is $\overline{\mathfrak{M}}(\mathfrak{R})$,

$$WFW^*E' = F'E'.$$

But it follows from Theorem 1.24 that $W^*E'=W^*$. Hence

$$WFW^* = F'E'.$$

But (S) Theorem 2.37 now yields, since WFW^* is a projection as shown above, $F'E'=E'F'$ and hence $WFW^*=E'F'$ and

$$(I - E)F' = F' - E'F' = F' - F'E' = F'(I - E').$$

Hence by (S) Theorem 2.37, $(I-E')F'$ is a projection and its range is included in $\mathfrak{H}_2\ominus\overline{\mathfrak{M}}(\mathfrak{R})=\mathfrak{R}^*$. Let $F_0'=(I-E')F'$; then

$$F' = E'F' + (I - E')F' = WFW^* + F_0'.$$

Conversely, since the range of WFW^* is included in the range of W, $\overline{\mathfrak{M}}(\mathfrak{R})$, if F_0' is a projection, whose range is included in $\mathfrak{H}_2\ominus\overline{\mathfrak{M}}(\mathfrak{R})=\mathfrak{R}^*$, then

$$F' = WFW^* + F_0'$$

is a projection by (S) Theorem 2.37 and

$$TF \supseteq WFW^*T = (WFW^* + F_0')T = F'T.$$

We have thus shown that if F breaks T, F' must be in the form stated and any such F' will satisfy the definition of breakage.

Theorem V. *If T is a closed linear transformation in $\mathfrak{H}$, with domain everywhere dense in $\mathfrak{H}$, then F reduces T if and only if it reduces B and W.*

Now if F reduces T, it breaks T, since we may take $F'=F$, and hence it reduces B and we obtain by the previous theorem

$$F = F' = WFW^* + F_0',$$

where F_0' is a projection, whose range is included in $\mathfrak{N}^*$, and hence $W^*F'=0$. Also in the proof of the previous theorem we have shown that $EFW^*=FW^*$ and so we obtain

$$W^*F = W^*(WFW^* + F_0') = W^*WFW^* + W^*F_0' = EFW^* + 0 = FW^*.$$

Since the domain of W^* is $\mathfrak{H}$, this yields that F reduces W^*.

But if F reduces W^*, it reduces W, since the equation $FW^*=W^*F$ implies by taking adjoints $W^{**}F=FW^{**}$ or $WF=FW$.

On the other hand if F reduces both B and W, it reduces T. For since the domain of B is $\mathfrak{D}$, the domain of T, and since F reduces B, $Ff\epsilon\mathfrak{D}$ for all $f\epsilon\mathfrak{D}$ and

$$TFf = WBFf = WFBf = FWBf = FTf.$$

Theorem VI. *Let T be as in Theorem* 1.24. *Then there exists a set of linear transformations* $\{T_i\}$, $i=0, \pm 1, \pm 2, \cdots$, *each closed and limited with a limited inverse, with domain $\mathfrak{D}_i$ and range $\mathfrak{R}_i$, such that $\mathfrak{D}_i \perp \mathfrak{D}_j$ and $\mathfrak{R}_i \perp \mathfrak{R}_j$ for for $i\neq j$ and such that for all $f\epsilon\mathfrak{H}_1$,*

$$f = Ff + \sum_{-\infty}^{\infty} f_i, \qquad f_i\epsilon\mathfrak{D}_i, \qquad Tf = \sum_{-\infty}^{\infty} T_if_i,\dagger$$

where F is the projection on the manifold of zeros of T.‡

We take any set of numbers $\{\alpha_i\}$ such that $i=0, \pm 1, \pm 2, \cdots, \alpha_i>\alpha_j$ if $i>j$, $\lim_{i\to-\infty}\alpha_i=0$, $\lim_{i\to\infty}\alpha_i=\infty$. We let $F_i=E(\alpha_i)-E(\alpha_{i-1})$, where $E(\lambda)$ is the resolution of the identity corresponding to B. Now it is shown in (S) Theorem 5.9 that the range of F_i is in the domain of B which is $\mathfrak{D}$, the domain of T. Hence we may take T_i as the contraction T, with domain the range of F_i.

It is also shown in the proof of (S) Theorem 5.7 that F_i is a projection and that the ranges of F_i and F_j for $i\neq j$ are mutually orthogonal or that $\mathfrak{D}_i \perp \mathfrak{D}_j$. In the proof of (S) Theorem 5.9 we learn that

$$\|BF_if\|^2 = \int_{\alpha_{i-1}}^{\alpha_i} \lambda^2 d\|E(\lambda)f\|^2. \tag{1}$$

† This equation is also to imply that Tf has a meaning if and only if the expression on the right has a meaning.

‡ Theorem VI compensates in the study of the inverses of contractions of T for the lack of compactness of Hilbert space. In treating certain types of partial differential equations, it has been possible to use spaces of the Banach type having a species of compactness. Cf. J. Schauder, Mathematische Annalen, vol. 106, pp. 661–721.

Hence we have that

$$\alpha_{i-1}^2(\|E(\alpha_i)f\|^2 - \|E(\alpha_{i-1})f\|^2) \leqq \|BF_if\|^2 \leqq \alpha_i^2(\|E(\alpha_i)f\|^2 - \|E(\alpha_{i-1})f\|^2).$$

Now if f is in the range of F_i, it is by (S) Theorem 2.37 perpendicular to the range of $E(\alpha_{i-1})$ and furthermore $E(\alpha_i)f=f$. Hence for all $f\epsilon\mathfrak{D}_i$,

$$\alpha_{i-1}^2\|f\|^2 \leqq \|B_if\|^2 \leqq \alpha_i^2\|f\|^2. \tag{2}$$

Since for $f\epsilon\mathfrak{D}_i$,

$$\|T_if\| = \|Tf\| = \|Bf\| = \|B_if\|,$$

(2) yields that T_i is limited with a limited inverse, and since its domain is a closed linear manifold, it is closed.

Now (S) Theorem 5.8 tells us that $\mathfrak{D}_i$ reduces B and hence by Theorem IV we have that F_i breaks T.

We will show that $\mathfrak{R}_i\perp\mathfrak{R}_j$, $i\neq j$. For let $T_if\epsilon\mathfrak{R}_i$ and $T_jg\epsilon\mathfrak{R}_j$. Hence $F_if=f$, $F_jg=g$. Now since T and B are isometrically equivalent, and F_i and F_j reduce B,

$$(T_if, T_jg)_2 = (B_if, B_jg)_1 = (B_iF_if, B_jF_jg)_1 = (F_iB_if, F_jB_jg) = 0,$$

since the ranges of F_i and F_j are mutually orthogonal.

By Definition 5.1 (S),

$$F + \sum_{-\infty}^{\infty} F_i = F + \lim_{\alpha_n\to\infty,\alpha_{-m}\to 0}(E(\alpha_n) - E(\alpha_{-m})) = F + I - E(0) = I - E(0-),$$

remembering that F is the manifold of zeros of B. Since B is not negative definite ((S) Definition 2.14), (S) Theorem 5.12 yields that $E(0-)=0$. Hence

$$F + \sum_{-\infty}^{\infty} F_i = I.$$

Hence for any $f\epsilon\mathfrak{H}_1$,

$$f = Ff + \sum_{-\infty}^{\infty} F_if = Ff + \sum_{-\infty}^{\infty} f_i,$$

letting $F_if=f_i$, $f_i\epsilon\mathfrak{D}_i$.

Since $BFf=0$, B is linear and $\lim_{i\to\infty}\alpha_i=0$,

$$\begin{aligned}\sum_{-\infty}^{N} B_if_i &= \sum_{-\infty}^{N} B(E(\alpha_i) - E(\alpha_{i-1}))f = BE(\alpha_n)f - \lim_{i\to-\infty} BE(\alpha_i)f\\ &= BE(\alpha_n)f - \lim_{\alpha_i\to 0} BE(\alpha_i)f = BE(\alpha_n)f - \lim_{\alpha_i\to 0} B(E(\alpha_i) - F)f.\end{aligned}$$

Since B is not negative definite, (S) Theorem 5.12 yields $F=E(0)$, and by Theorem 5.9 (S),

$$\|B(E(\alpha_i) - E(0))f\|^2 = \int_0^{\alpha_i} \lambda^2 d\|E(\lambda)f\|^2 \leqq \alpha_i^2\|f\|^2.$$

Hence

$$\lim_{\alpha_i \to 0} B(E(\alpha_i) - F)f = \lim_{\alpha_i \to 0} B(E(\alpha_i) - E(0))f = 0,$$

and from (3) we get

$$\sum_{-\infty}^{N} B_i f_i = B(E(\alpha_n)f).$$

Since B is not negative definite, (S) Theorem 5.12 yields $E(-N)=0$ for $N>0$. Let $\Delta_N=(-N, \alpha_N)$; then

$$\begin{aligned}\sum_{-\infty}^{\infty} B_i f_i &= \lim_{N\to\infty} \sum_{-\infty}^{N} B_i f_i = \lim_{N\to\infty}\left(\sum_{-\infty}^{N} B_i f_i\right) - \lim_{N\to\infty} BE(-N)f \\ &= \lim_{N\to\infty} B(E(\alpha_n) - E(-N))f = \lim_{N\to\infty} BE(\Delta_N)f.\end{aligned}$$

Now $\lim_{n\to\infty} \alpha_n = \infty$, and in the proof of (S) Theorem 5.9, it is shown that the limit on the right exists if and only if Bf exists, and if Bf exists, then this limit equals Bf. Hence

$$Bf = \sum_{-\infty}^{\infty} B_i f_i.$$

Now since W is isometric on $\overline{\mathfrak{M}}(\mathfrak{R}_b)$,

$$\sum_{-\infty}^{\infty} WB_i f_i$$

converges if and only if $\sum_{-\infty}^{\infty} B_i f_i$ converges, and hence since the domain of T is the same as the domain of B if and only if $f\epsilon\mathfrak{D}$, and when $f\epsilon\mathfrak{D}$, we have

$$Tf = WBf = W\sum_{-\infty}^{\infty} B_i f_i = \sum_{-\infty}^{\infty} WB_i f_i = \sum_{-\infty}^{\infty} T_i f_i,$$

since W is continuous.

THEOREM VII. *Let T be a closed and linear transformation from $\mathfrak{H}_1$ to $\mathfrak{H}_2$, with domain everywhere dense in $\mathfrak{H}_1$. T^* exists and is a closed linear transformation from $\mathfrak{H}_2$ to $\mathfrak{H}_1$, with domain everywhere dense in $\mathfrak{H}_2$ (Theorem 1.13). If $\mathfrak{M}_1\epsilon\mathfrak{H}_1$ and $\mathfrak{M}_2\epsilon\mathfrak{H}_2$ are the ranges of two projections F_1 and F_2 and are such that*

(1) *if $f\epsilon\mathfrak{D}$, then $F_1 f\epsilon\mathfrak{D}$;*
(2) *if $f\epsilon\mathfrak{D}^*$, then $F_2 f\epsilon\mathfrak{D}^*$;*
(3) *if $f\epsilon\mathfrak{M}_1\cdot\mathfrak{D}$, then $Tf\epsilon\mathfrak{M}_2$;*
(4) *if $f\epsilon\mathfrak{M}_2\cdot\mathfrak{D}^*$, then $T^*f\epsilon\mathfrak{M}_1$;*

then F_1 breaks T and $TF_1f=F_2Tf$ for all $f\epsilon\mathfrak{D}$.

Due to condition (1), if we can prove that $TF_1f=F_2Tf$ for every $f\epsilon\mathfrak{D}$, the theorem follows. From condition (3) we notice that for all $f\epsilon\mathfrak{D}$,

$$F_2TF_1f = TF_1f.$$

Now if $f\epsilon\mathfrak{D}$, and $g\epsilon\mathfrak{D}^*$, we have by condition (2)

$$\begin{aligned}(TF_1f - F_2Tf,\, g)_2 &= (F_2TF_1f - F_2Tf,\, g)_2 = (TF_1f - Tf,\, F_2g)_2 \\ &= (F_1f - f,\, T^*F_2g)_1.\end{aligned}$$

But $F_1f-f\epsilon\mathfrak{H}_1\ominus\mathfrak{M}_1$, while by (4) $T^*Fg\epsilon\mathfrak{M}_1$. Hence for all $f\epsilon\mathfrak{D}$, $g\epsilon\mathfrak{D}^*$,

$$((TF_1 - F_1T)f,\, g)_2 = 0.$$

But $\mathfrak{D}^*$ is everywhere dense in $\mathfrak{H}_2$, hence for every $f\epsilon\mathfrak{D}$,

$$TF_1f = F_2Tf.$$

6. Functions of class H

Let S be a bounded connected open region. Let S_x denote the projection of S on the x-axis, S_y denote the projection of S on the y-axis. We shall use the symbol $m_1(E)$ to denote the linear measure of a linear set.

Definition II. *A function $f=f(x, y)$ shall be said to be of class H on S (symbolically $f\epsilon H_s$), if it satisfies the following conditions:*

1. *It has a summable square and has almost everywhere first and second partial derivatives of summable square on the region S.*

2.a. *There exists a set Γ_y on the y-axis, such that $\Gamma_y\subseteq S_y$, $m_1(S_y-\Gamma_y)=0$ and if $y\epsilon\Gamma_y$, then $f(x, y)$, $f_x(x, y)$ and $f_y(x, y)$ are absolutely continuous functions of x, defined for all x such that $(x, y)\epsilon S$.*

2.b. *There exists a set Γ_x on the x-axis, such that $\Gamma_x\subseteq S_x$, $m_1(S_x-\Gamma_x)=0$, and if $x\epsilon\Gamma_x$, then $f(x, y)$, $f_x(x, y)$ and $f_y(x, y)$ are absolutely continuous functions of y defined for all y such that $(x, y)\epsilon S$.*

For f and $g\epsilon H_s$, let

$$\begin{aligned}\|f - g\|^2 = \iint_S (&|f - g|^2 + |f_x - g_x|^2 + |f_y - g_y|^2 + |f_{xx} - g_{xx}|^2 \\ &+ |f_{xy} - g_{xy}|^2 + |f_{yx} - g_{yx}|^2 + |f_{yy} + g_{yy}|^2)dS.\end{aligned}$$

Lemma. *Given a sequence of functions $\{f^{(n)}\}$, $f^{(n)}\epsilon H_s$, such that*

$$\lim_{n+m\to\infty} \|f^{(n)} - f^{(m)}\| = 0.$$

Then there exists an $f\epsilon H_s$, such that

$$\lim_{n\to\infty} \|f - f^{(n)}\| = 0.$$

The sequence $\{f_{x,x}^{(n)}\}$ is convergent in the mean, hence there exists, by the Riesz-Fischer Theorem, a function $f_{1,1}$ of summable square such that

$$\lim_{n\to\infty}\int\int_S |f_{x,x}^{(n)} - f_{1,1}|^2 dS = 0.$$

But by Fubini's Theorem, we have that

$$\int\int_S |f_{x,x}^{(n)} - f_{1,1}|^2 dS = \int_{S_y} dy \int_{S_1(y)} |f_{x,x}^{(n)} - f_{1,1}|^2 dx,$$

where $S_1(y)$ denotes the set of points in S whose ordinate is y. Now if we set

$$\phi_n^2(y) = \int_{S(y)} |f_{x,x}^{(n)}(x, y) - f_{1,1}(x, y)|^2 dx,$$

we see from the above that

$$\lim_{n\to\infty}\int_{S_y} |\phi_n(y) - 0|^2 dy = 0.$$

Hence by the Riesz-Fischer Theorem, there exists a subsequence $\{\psi_n\}$ of $\{\phi_n\}$ (and a corresponding subsequence $\{g^{(n)}\}$ of the sequence $\{f^{(n)}\}$), and a linear set $C'_{x,x}$, $C'_{x,x}\subseteq S_y$ for which $m_1(S_y - C'_{xx}) = 0$, such that if $y\epsilon C'_{x,x}$, then

$$(1)\qquad 0 = \lim_{n\to\infty}\psi_n^2(y) = \lim_{n\to\infty}\int_{S(y)} |g_{x,x}^{(n)}(x, y) - f_{1,1}(x, y)|^2 dx.$$

Now the sequence $\{g_{x,y}^{(n)}\}$ converges in the mean also, to some function of summable square $f_{1,2}$, and in a similar manner, we can find a subsequence $\{h^{(n)}\}$ of the sequence $\{g^{(n)}\}$ and a set $C'_{x,y}\subseteq S_x$, $m_1(S_x - C'_{,y}) = 0$, such that for $x\epsilon C'_{x,y}$,

$$(2)\qquad \lim_{n\to\infty}\int_{S_2(x)} |h_{x,y}^{(n)}(x, y) - f_{1,2}(x, y)|^2 dy = 0.$$

The sequence $\{h_x^{(n)}\}$ converges in the mean. Hence by the Riesz-Fischer Theorem, there exists a subsequence $\{k_x^{(n)}\}$, a function f_1, and a set $S'\subseteq S$, $mS' = mS$, such that on S'

$$(3)\qquad \lim_{n\to\infty} k_x^{(n)} = f_1.$$

Now let $\Gamma_y^{(n)}$ correspond to $f^{(n)}$ as Γ_y did to f in the definition of a function of class H on S. Let $C_{x,x}$ denote $C_{x,x}\Pi\Gamma_y^{(n)}$. One can readily verify that

$m_1(S_y - C_{x,x}) = 0$. Similarly, we define $C_{x,x} = C'_{x,y} \Pi \Gamma_x^{(n)}$ and obtain

$$m_1(S_x - C_{x,y}) = 0.$$

Now since $m(S'(C_{x,y} \times C_{x,x})) = mS \neq 0$, there is a point P: $(a, b) \epsilon S'(C_{x,y} \times C_{x,x})$. Let S_1 consist of all points $(x, y) \epsilon S$ such that either $y \epsilon C_{x,x}$ or $x \epsilon C_{x,y}$. Let $Q:(x_0, y_0) \epsilon S_1$. Then we can connect P and Q by a polygonal line Γ, $PR_1R_2 \cdots R_nQ$, such that each segment $PR_1 \subset S_1$, $R_iR_{i+1} \subset S_1$, $R_nQ \subset S_1$ and each segment is parallel to an axis.

Now if a segment of Γ is parallel to the x-axis, it follows from (1) that $f_{1,1}$ is measurable (linearly) along the segment and has a summable square, by the Riesz-Fischer Theorem. Hence $f_{1,1}$ is also summable since for a finite interval the summability of the square implies the summability of the function. Similarly along a segment of Γ parallel to the y-axis, $f_{1,2}$ is summable by (2). Thus we may define

$$f_x(Q) = f_1(P) + \int_\Gamma f_{1,1}dx + f_{1,2}dy. \tag{4}$$

We now show that the definition of $f_x(Q)$ is independent of the choice of Γ or of P, in fact that $f_x(Q) = f_1(Q)$, whenever the former is defined.

For, suppose R_iR_{i+1} is a segment of Γ, parallel to the x-axis, i.e., R_i is (x_1, y), R_{i+1} is (x_2, y), $y \epsilon C_{x,x}$. Since $C_{x,x} \supseteq \Gamma_x^{(n)}$, $f_{x,y}(x, y)$ is absolutely continuous, and

$$\int_{x_2}^{x_1} f_{x,x}^{(n)}(\xi, y)d\xi = f_x^{(n)}(x_1, y) - f_x^{(n)}(x_2, y), \tag{5}$$

and by Schwarz's Inequality (remembering that $\{k^{(n)}\}$ is a subset of $\{f^{(n)}\}$),

$$\begin{aligned} &| k_x^{(n)}(x_1, y) - f_x(x_1, y) - (k_x^{(n)}(x_2, y) - f_x(x_2, y)) |^2 \\ &= \left| \int_{x_2}^{x_1} (k_x^{(n)}(\xi, y) - f_{1,1}(\xi, y))d\xi \right|^2 \\ &\leqq | x_1 - x_2 | \left| \int_{x_2}^{x_1} | k_{x,x}^{(n)}(\xi, y) - f_{1,1}(\xi, y) |^2 d\xi \right|. \end{aligned}$$

Since $\{k^{(n)}\} \subseteq \{g^{(n)}\}$, it follows from (1) that

$$\lim_{n\to\infty} | (k_x^{(n)}(x_1, y) - f_x(x_1, y)) - (k_x^{n)}(x_2, y) - f_x(x_2, y)) |^2 = 0. \tag{6}$$

If, on the other hand, R_jR_{j+1} is parallel to the y-axis and if $R_j = (x, y_1)$, $R_{j+1} = (x, y_2)$, $x \epsilon C_{x,y}$, we obtain similarly

$$\lim_{n\to\infty} | (k_x^{(n)}(x, y_1) - f_x(x, y_1)) - (k_x^{(n)}(x, y_2) - f_x(x, y_2)) |^2 = 0. \tag{7}$$

Now either (6) or (7) applies to PR_1. We also have by the choice of $P:(a, b)$

$$\lim_{n\to\infty} k_x^{(n)}(a, b) = f_x(a, b).$$

Hence

$$\lim_{n\to\infty} k_x^{(n)}(R_1) = f_x(R_1).$$

Similarly, we may continue to apply either (6) or (7), obtaining

$$\lim_{n\to\infty} k_x^{(n)}(R_i) = f_x(R_i), \quad \lim_{n\to\infty} k_x^{(n)}(Q) = f_x(Q).$$

But the limit of the sequence is independent of both P and the choice of Γ. We have also shown that

$$f_x(Q) = f_1(Q)$$

wherever the former is defined.

Recalling (3), that $mS_1 = mS$, and the Riesz-Fischer Theorem, we obtain that

$$\begin{aligned} 0 &= \lim_{n\to\infty} \iint_S |k_x^{(n)} - f_1|^2 dS = \lim_{n\to\infty} \iint_{S_1} |k_x^{(n)} - f_1|^2 dS \\ &= \lim_{n\to\infty} \iint_{S_1} |k_x^{(n)} - f_x|^2 dS = \lim_{n\to\infty} \iint_S |k_x^{(n)} - f_x|^2 dS, \end{aligned}$$

as well as the fact that f_x has a summable square. These two facts readily yield, by Schwarz's Inequality, that

$$\begin{aligned} 0 &\leqq \liminf_{n\to\infty} \left(\iint_S |f_x^{(n)} - f_x|^2 dS \right)^{1/2} \\ &\leqq \limsup_{n\to\infty} \lim_{m\to\infty} \left(\iint_S |f_x^{(n)} - k_x^{(m)} + k_x^{(m)} - f_x|^2 dS \right)^{1/2} \\ &\leqq \limsup_{n\to\infty} \limsup_{m\to\infty} \left(\iint_S |f_x^{(n)} - k_x^{(m)}|^2 dS \right)^{1/2} + \left(\iint_S |k_x^{(m)} - f_x|^2 dS \right)^{1/2} \\ &= \lim_{n,m\to\infty} \left(\iint_S |f_x^{(n)} - k_x^{(m)}|^2 dS \right)^{1/2} + 0 = 0. \end{aligned}$$

In a similar fashion we may establish the existence of two sets S_2, S_0, $S_2 \subseteq S$, $S_0 \subseteq S$, $mS_2 = mS_0 = mS$, and of two functions f_y and f, such that

$$f_y(Q) = f_2(P_2) + \int_\Gamma f_{2,1}dx + f_{2,2}dy, \quad \lim_{n\to\infty} \iint_S |f_y^{(n)} - f_y|^2 dS = 0,$$

and

$$f \equiv f(P_3) + \int_\Gamma f_1 dx + f_2 dy, \quad \lim_{n\to\infty} \int\int_S |f^{(n)} - f|^2 dS = 0.$$

Now if a sequence of functions, defined on a finite linear interval, converges in the mean, and if a second sequence of functions, each of which is an indefinite integral of the corresponding function in the first set, is such that it converges at a single point, the second sequence converges in the mean.

This proves that C_x' can be taken in such a manner that $C_x' \supseteq C_{x,x}$ when one considers the defining equations above and their analogue in the proof of the existence of f. Hence $C_x \subseteq C_{x,x}$. In a similar manner, we take C_y' in such a way that $C_y' \supseteq C_{y,y}$ and obtain that $C_y \supseteq C_{y,y}$.

We also have, since for $y=c$, $c \epsilon C_{x,x}$, $f_1 = f_x$ is continuous, that

$$\frac{\partial f}{\partial x}(x, c) = f_1 = f_x.$$

Similarly, for a point whose abscissa is in $C_{y,y}$,

$$\frac{\partial f}{\partial y} = f_2 = f_y.$$

Finally we see that, on each line of S_1, almost everywhere

$$\frac{\partial^2 f}{\partial x^2} = f_{1,1}, \quad \frac{\partial^2 f}{\partial x \partial y} = f_{1,2}.$$

These results yield, upon inspection, that f is of class H on S and that

$$\lim_{n\to\infty} \|f^{(n)} - f\| = 0$$

irrespective of any further extension of the definition of f.

7. The space $\mathfrak{W}$

Let us define the inner product of two functions of class H on S as follows:

$$(u, v) = \int\int_S (u\bar{v} + u_x\bar{v}_x + u_y\bar{v}_y + u_{x,x}\bar{v}_{x,x} + u_{x,y}\bar{v}_{x,y} + u_{y,x}\bar{v}_{y,x} + u_{y,y}\bar{v}_{y,y})dS.$$

We wish to set up a Hilbert space by means of functions of class H on S and the above inner product just as this is done in the classical theory, with $\mathfrak{L}_2$ and $\int\int u\bar{v}dS$. As is the case of $\mathfrak{L}_2$, we are forced to take as our points classes of equivalent functions.

We set up the classes by means of the following statement: Two functions u and v belong to the same class if and only if

$$\|u - v\| = 0.$$

We define the sum of two classes as the class to which the sum of an element of the first class and an element of the second class belongs. The product of a class and a number c is the class to which c times a function of the class belongs and the inner product of two classes f and g is defined as

$$(f, g) = (u, v)$$

where u and v belong to f and g respectively. One can readily show by familiar methods that these definitions yield a unique result in each case.

THEOREM VIII.* *The space $\mathfrak{W}$, whose elements are the classes defined above with the above definition of addition, scalar multiplication, and inner product, is a Hilbert space.*†

The completeness of $\mathfrak{W}$ is the lemma of the preceding section. The remainder of the proof is immediate and will be omitted.

Turning now to the consideration of equation (A) for functions u of class H, we have, by Schwarz's Inequality and the boundedness of the coefficients,

$$(\alpha)\quad \iint |L(u)|^2 dS \leqq \iint_S (|A_{20}|^2 + 2|A_{1,1}|^2 + \cdots + |A_{0,0}|^2)$$
$$\cdot(|u_{xx}|^2 + |u_{x,y}|^2 + \cdots + |u|^2)dS \leqq C^2\|u\|^2,$$

where C is a constant.

Thus we see that the functions $L(u)$, $u \epsilon H_s$, have a summable square on S. We can set up the usual space $\mathfrak{L}_2$ (cf. (S) Theorem 1.24) and corresponding to the set of $L(u)$'s we have a subset of $\mathfrak{L}_2$. Let $\mathfrak{R}$ be the closed linear manifold determined by this subset. We suppose that $\mathfrak{R}$ is of infinite dimensionality and thus is a Hilbert space (cf. (S) Theorem 1.18).

The equation (A) determines a transformation T from $\mathfrak{W}$ to $\mathfrak{R}$ defined as follows: If $f \epsilon \mathfrak{W}$ and u is one of its functional representatives, i.e., $u \epsilon f$, then

* For other examples of spaces proposed for the study of partial differential equations, there is that of O. Nikodym (Journal de Mathématiques, vol. 9 (1933), pp. 95–109) who uses as his scalar product

$$(f, g) = \iint_S (f_x g_x + f_y g_y + f_s \bar{g}_s) dS$$

and obtains a simple proof of a theorem of M. Zaremba.

Spaces of the Banach type have been effectively used by J. Schauder (Mathematische Annalen, vol. 106 (1932), pp. 661–721) in the study of elliptic partial differential equations. Hyperbolic differential equations of a special form with initial conditions have been treated by D. C. Lewis (these Transactions, vol. 35, pp. 792–823) by methods which involve $\mathfrak{L}_p$.

The work of Ritz can be regarded as having to do with spaces in which the length of a function u is the integral J which he minimizes.

† (S) Definition 1.1.

from the above we see that $L(u)$ is a member of a class g, which is a point in $\mathfrak{K}$, and we define $Tf=g$. Thus it follows from (α) that T is a limited transformation with domain $\mathfrak{W}$ and hence T is closed. The range of T is everywhere dense in $\mathfrak{K}$, by virtue of the definition of $\mathfrak{K}$.

8. Application to differential operators

We are now in a position to apply our results on transformations between Hilbert spaces to the transformation represented by equation (A). Specifically we give the significance in terms of functions of Theorems 1.16 and II, which we may apply since we have restricted our range space to the closed linear manifold $\mathfrak{K}$, in $\mathfrak{L}_2$, determined by the range of T.

The following notation is useful.

Definition III. *$w=_m u$ is to mean that $w=u$ almost everywhere on S. If $f_i \epsilon H_s$ and if the sequence $\{f_i\}$ is such that*

$$\lim_{m+n\to\infty}\left\| \sum_1^n f_i - \sum_1^m f_i \right\| = 0,$$

then $\sum_1^\infty f_i$ is to denote the function whose existence is proved by the lemma of §6.

$u' =_{m+} u''$ is to mean $\|u'-u''\|=0$.

We then obtain

Theorem 1.16′. *Given the equation* (A) *of the introduction, and a bounded connected region S. There exist two orthonormal sets of functions $\{u_i\}$ and $\{v_i\}$, u_i and $v_i \epsilon H_s$, such that the corresponding elements in $\mathfrak{W}$ are mutually orthogonal and are such that*

(1) *a necessary and sufficient condition that $u \epsilon H_s$ be such that $L(u)=_m 0$, is that*

$$u =_{m+} \sum_1^\infty (u, u_i)u_i, \quad \sum_1^\infty |(u, u_i)|^2 < \infty;$$

(2) *if $v =_{m+} \sum_1^\infty (v, v_i)v_i$ and $L(v) =_m 0$, then $v =_{m+} 0$;*

(3) *if $w \epsilon H_s$, then*

$$w =_{m+} \sum_1^\infty (w, u_i)u_i + \sum_1^\infty (w, v_i)v_i, \quad \sum_1^\infty |(w, u_i)|^2 + \sum_1^\infty |(w, v_i)|^2 < \infty .$$

This is of course Theorem 1.16 applied to this special case and then using the fact that in any closed linear manifold there exists an orthonormal set complete in the closed linear manifold, for $\mathfrak{N}$ and $\mathfrak{H}\ominus\mathfrak{N}$ ((S) Theorems 1.19 and 1.14). If we have given a complete orthonormal set $\{f_i\}$ in $\mathfrak{W}$ and a

complete orthonormal set $\{g_i\}$ in $\mathfrak{K}$ we can actually exhibit the sets u_i and v_i. For

$$T^*g_i = \sum_{j=1}^{\infty}(T^*g_i, f_j)_w f_j = \sum_{1}^{\infty}(g_i, Tf_j)_k f_j$$

and $\{T^*g_i\}$ determine $\mathfrak{H}\ominus\mathfrak{N}$ by Theorems II and 1.16, and by the method in the note on construction, we can obtain a set which determines $\mathfrak{N}$.

We define $L^*(v)$ for any v of a class g, which is an element of $\mathfrak{K}$, as a function of class H which is a member of T^*g.

THEOREM II′. *Let $\{\psi_i\}$ denote the set obtained by orthonormalizing the set $\{L^*(L(v_i))\}$ where the v_i are as in Theorem* 1.16′. *Then $\{\psi_i\}$ determines the the same closed linear manifold as the set $\{v_i\}$ of Theorem* 1.16′. *For each ψ_i there is a χ_i such that $L^*(\chi_i) =_{m+} \psi_i$. A necessary and sufficient condition that a function w, a member of an element g of $\mathfrak{K}$, be such that there is a $u \epsilon H$, for which $L(u) =_m w$, is that*

$$\sum_{1}^{\infty} |(w, \chi_i)|^2 < \infty;$$

when w satisfies this condition and

$$v =_{m+} \sum_{1}^{\infty}(w, \chi_i)\psi_i,$$

then $L(v) =_m w$.

This is just the application of Theorem II to the special case treated here.

Theorem II′ offers a generalization of the methods of Ritz† in the sense that it generalizes the formulas upon which actual calculations would be based.

We have restricted the coefficients to be bounded and upon this assumption the limitedness of this transformation was shown. In the unbounded cases, some knowledge of the resolution of the identity corresponding to T^*T must be had before results analogous to those of Theorems 1.16′ and II′ can be obtained.

The following examples indicate what results are to be expected in applying the above methods. Let us suppose S is a square of side 2π with sides parallel to an axis. We will show that the range of the transformation associated with

$$L(u) = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = v$$

is the whole space $\mathfrak{L}_2$. Let v be a function of summable square; then

† Loc. cit.

$$v = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} a_{n,m} \frac{e^{inx+imy}}{2\pi}; \quad a_{n,m} = \frac{1}{2\pi} \int\int_S e^{-inx-imy} v dS.$$

Now when we orthonormalize the sequence $\{e^{inx+imy}\}$ in $\mathfrak{W}$ we obtain

$$\left\{ \frac{1}{2\pi} \left(1 + m^2 + n^2 + (m^2 + n^2)^2 \right)^{-1/2} e^{inx+imy} \right\}.$$

Now since $\sum_{-\infty}^{\infty} \sum_{-\infty}^{\infty} |a_{n,m}|^2 < \infty$,

$$(1) \qquad \sum_{n^2+m^2>0} \sum |a_{n,m}|^2 \frac{1 + n^2 + m^2 + (m^2 + n^2)^2}{(m^2 + n^2)^2} < \infty,$$

let

$$u = \sum_{n^2+m^2>0} \sum a_{n,m} \frac{(1 + n^2 + m^2 + (m^2 + n^2)^2)^{1/2}}{m^2 + n^2} \frac{1}{2\pi} \frac{e^{inx+imy}}{(1 + n^2 + m^2 + (n^2 + m^2)^2)^{1/2}}$$

$$= \sum_{n^2+m^2>0} \sum \frac{a_{n,m}}{2\pi} \frac{e^{inx+imy}}{m^2 + n^2}.$$

From (1), we see that $u \epsilon H_s$ and $a_{00}x^2/(4\pi) + u \epsilon H_s$ and hence

$$L_1\left(\frac{a_{00}x^2}{4\pi^2} + u \right) =_m v.$$

Thus when T_1 is restricted to $\mathfrak{W} \ominus \mathfrak{N}$ so that it has an inverse, Theorem 1.25 yields that this inverse is bounded.

However, for the same square in the case of

$$L_2(u) = \frac{\partial^2 u}{\partial x \partial y} = v,$$

while the range of T is everywhere dense in $\mathfrak{L}_2$, it is not $\mathfrak{L}_2$. For we have no difficulty in finding a $u_{n,m}$ such that

$$\frac{\partial^2 u_{n,m}}{\partial x \partial y} = e^{inx+imy},$$

thus showing that the range determines $\mathfrak{L}_2$ and since it is a linear manifold that it is everywhere dense in $\mathfrak{L}_2$.

However there is no $u \epsilon \mathfrak{W}$ such that

$$(2) \qquad \frac{\partial^2 u}{\partial x \partial y} = \log x,$$

although $\log x \epsilon \mathfrak{L}_2$. For if u satisfies (2) it is in the form

$$u = xy(\log x - 1) + X(x) + Y(y),$$

and one can easily show that $\partial^2 u/\partial x^2$ does not have a summable square.

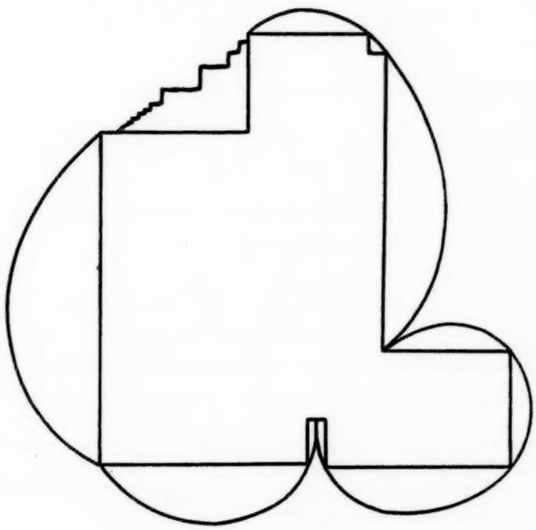
9. Boundary values of f_x and f_y

In the remaining sections, we give a method for the solution of boundary value problems, specifically the problem formulated in §10. We must first restrict S so as to have a boundary which fulfils the following conditions.

Let the boundary B of S be a rectifiable curve such that a polygon $P_1P_2 \cdots P_n$, $P_i=(x_i, y_i)$, with P_iP_{i+1} parallel to an axis and the side of a rectangle whose interior is in S and with the interior of $P_1P_2 \cdots P_n$ in S, can be constructed with certain vertices on B in such a manner that the arc of B between two vertices which are successive along the curve is such that either

(1) *it is intercepted by two alternate vertices of the polygon P_i and P_{i+2} and if (x, y) is a point of the arc P_iP_{i+2}, there is a relationship $y=Y'(x)$, which is such that Y' is monotonic and single-valued except possibly at a denumerable set of points, where the continuum of values between the limits on the right and left are assumed, or*

(2) *it is intercepted by two successive vertices P_j and P_{j+1} and* (a) *$x_j=x_{j+1}$ and the arc P_jP_{j+1} is given by an equation $x=X(y)$, $y_j \leqq y \leqq y_{j+1}$, where $X(y)$ is single-valued and has bounded Dini derivatives or* (b) *$y_{j+1}=y_j$ and the arc P_jP_{j+1} is given by the equation $y=Y(x)$ where Y is single-valued and has bounded Dini derivatives.* (See the figure.)



As to the relationship between conditions (1) and (2) it may be pointed out that a monotonic curve may be rotated into one with bounded Dini derivatives. It should also be pointed out that these restrictions rule out comparatively smooth curves with outward pointing cusps.*

We shall suppose from now on that S satisfies the above conditions and that $f(x, y)\epsilon H_s$.

LEMMA I. *Let P_iP_{i+1} be a segment of the polygon described above. Let us suppose that $y_i = y_{i+1}$ and $x_i < x_{i+1}$; then $\lim_{y\to y_i} f_x(x, y)$ exists for almost every x such that $x_i \leqq x \leqq x_{i+1}$. If $x\epsilon\Gamma_x$ and $(x, y_i)\epsilon S$, then $\lim_{y\to y_i} f_x(x, y) = f_x(x, y_i)$. $\lim_{y\to y_i} f_x(x, y)$ is measurable and has summable square on the interval $x_i \leqq x \leqq x_{i+1}$. There exists a constant K_i such that*

$$\int_{x_i}^{x_{i+1}} \left| \lim_{y\to y_i} f_x(x, y) \right|^2 dx \leqq K_i \|f\|^2.$$

Now by the specifications on the polygon $P_1P_2 \cdots P_n$ it is possible to find a rectangle $\Delta = P_iP_{i+1}AB$, $A = (x_{i+1}, a)$, $B = (x_i, a)$, whose interior is in S. Let us suppose $a < y_i$. Now

$$\iint_\Delta |f_x|^2 dS \tag{1}$$

exists. By Fubini's Theorem,

$$\int_a^{y_i} \left(\int_{x_i}^{x_{i+1}} |f_x(x, y)|^2 dx \right) dy \tag{2}$$

exists. Therefore there is a y', $a \leqq y' < y$, such that

$$\int_{x_i}^{x_{i+1}} |f(x, y')|^2 dx \leqq \frac{1}{y_i - a} \int_a^{y_i} \int_{x_i}^{x_{i+1}} |f_x(x, y)|^2 dx dy \leqq K^{(1)} \|f\|^2. \tag{3}$$

Now let Δ' be the rectangle with vertices P_i, P_{i+1}, (x_{i+1}, y'), (x_i, y'). Now $\Delta' \subseteq \Delta$ and we have by Fubini's Theorem

$$\iint_{\Delta'} |f_{x,y}|^2 dS = \int_{x_i}^{x_{i+1}} \int_{y'}^{y_i} |f_{x,y}|^2 dy\, dx, \tag{4}$$

and since the summability of the square on a bounded set implies summability, we have

* However, for our purposes, some outward pointing cusps must be excluded as the following example shows. Let S be the region of points (x, y) such that $0 < x < 1$, $x^3 > y > 0$. One can easily verify that $u^{1/2}\epsilon H_s$, and also that the values of u_x on the boundary do not have a summable square with respect to integration along the arc-length.

$$\iint_{\Delta'} f_{x,y}dS \tag{5}$$

exists. On applying Fubini's Theorem, we have that for almost every x, $x_i \leqq x \leqq x_{i+1}$, the integral

$$\int_{y'}^{y_i} f_{x,y}(x, y)dy \tag{6}$$

exists and is a measurable function of x for $x_i \leqq x \leqq x_{i+1}$. Now

$$\left| \int_{y'}^{y_i} f_x(x, y)dy \right|^2 \leqq (y_i - y') \int_{y'}^{y_i} | f_{x,y}(x, y) |^2 dy. \tag{7}$$

The existence of the integral on the right of (4) now implies that the function of x on the left of (7) is summable and hence the function (6) has a summable square and

$$\begin{aligned} &\int_{x_i}^{x_{i+1}} \left| \int_{y'}^{y_i} f_{x,y}(x, y)dy \right|^2 dx \\ &\leqq | y_i - y' | \int_{x_i}^{x_{i+1}} \int_{y'}^{y_i} | f_{x,y}(x, y) |^2 dydx \\ &= | y_i - y' | \iint_{\Delta'} | f_{x,y} |^2 dS \leqq K' \| f \|^2. \end{aligned} \tag{8}$$

Now for $x \epsilon \Gamma_x$ and for which (6) is defined and hence for almost every x such that $x_i \leqq x \leqq x_{i+1}$, we have

$$\begin{aligned} \lim_{y \to y_i} f_x(x, y) &= \lim_{y \to y_i} \left(f_x(x, y') + \int_{y'}^{y} f_{x,y}(x, \xi)d\xi \right) \\ &= f_x(x, y') + \int_{y'}^{y_i} f_{x,y}(x, \xi)d\xi. \end{aligned}$$

Hence $\lim_{y \to y_i} f_x(x, y)$ exists for almost every x such that $x_i \leqq x \leqq x_{i+1}$ and is a measurable function with a summable square on the interval $x_i \leqq x \leqq x_{i+1}$, since it is the sum of two functions which are measurable and have summable squares. Then from (4) and (8), by using the triangle theorem, we obtain

$$\int_{x_i}^{x_{i+1}} | \lim_{y \to y_i} f_x(x, y) |^2 dx \leqq K_i' \| f \|^2.$$

LEMMA II. *Let* arc P_iP_{i+2} *be an arc of the curve* B *such that case* (1) *of the specifications of* B *applies. Let us suppose that* $x_i < x_{i+1} = x_{i+2}$ *and* $y_i = y_{i+1} < y_{i+2}$. *Let* $Y(x)$ *be defined as the least of the numbers* $Y'(x)$ *for* x *fixed.** *Then obviously* $Y(x)$ *is a monotonically increasing single-valued function of* x, *and if* $f_x(x, Y(x))$ *is defined as the* $\lim_{y\to Y(x)} f_x(x, y)$, *when the latter exists, then* $f_x(x, Y(x))$ *is defined almost everywhere on the interval* $x_i \leqq x \leqq x_{i+2}$ *and is a measurable function of* x *with a summable square on this interval, and there is a* K *such that*

$$\int_{x_i}^{x_{i+2}} |f_x(x, Y(x))|^2 dx \leqq K\|f\|^2.$$

Let us denote the rectangle having P_iP_{i+1} and $P_{i+1}P_{i+2}$ among its sides as Δ. Then

$$\iint_{\Delta\cdot S} |f_{x,y}|^2 dS \tag{1}$$

exists. Hence by Fubini's Theorem

$$\int_{y_i}^{Y(x)} |f_{x,y}(x, y)|^2 dy \tag{2}$$

exists for almost every x such that $x_i \leqq x \leqq x_{i+2}$ and is a summable function of x such that

$$\int_{x_i}^{x_{i+1}} dx \int_{y_i}^{Y(x)} |f_{x,y}(x, y)|^2 dy = \iint_{\Delta\cdot S} |f_{x,y}|^2 dy. \tag{3}$$

Now since the summability of the square implies summability, the existence of (1) implies the existence of

$$\iint_{\Delta\cdot S} f_{x,y} dS \tag{4}$$

and by Fubini's Theorem

$$\int_{y_i}^{Y(x)} f_{x,y}(x, y) dx \tag{5}$$

exists for almost every x such that $x_i \leqq x \leqq x_{i+2}$ and is a measurable function on this interval.

Furthermore by Schwarz's Inequality when (2) and (5) exist, which is for almost every x in the interval $x_i \leqq x \leqq x_{i+2}$,

---

* It is to be remembered that $Y'(x)$ is not necessarily single-valued.

$$\left| \int_{y_i}^{Y(x)} f_{x,y}(x, y)dy \right|^2 \leqq | Y(x) - y_i | \int_{y_i}^{Y(x)} | f_{x,y}(x, y) |^2 dy$$
$$\leqq | y_{i+2} - y_i | \int_{y_i}^{Y(x)} | f_{x,y}(x, y) |^2 dy.$$

Since the function on the right is summable, the function on the left is also, and furthermore integrating and using (3) we obtain

$$\int_{x_i}^{x_{i+2}} \left| \int_{y_i}^{Y(x)} f_{x,y}(x, y)dy \right|^2 \leqq | y_{i+2} - y_i | \iint_{\Delta \cdot S} | f_{x,y} |^2 dS$$
$$\leqq | y_{i+2} - y_i | \, \|f\|^2. \tag{6}$$

Remembering that for $Y(x) = y_i$, by Lemma I, $f_x(x, Y(x))$ exists for almost every such x, we obtain for all x such that $x \epsilon \Gamma_x$ and for which (5) has been shown to hold and hence for almost every x such that $x_i \leqq x \leqq x_{i+2}$,

$$\begin{aligned} f_x(x, Y(x)) &= \lim_{y \to Y(x)^-} f_x(x, y) \\ &= \lim_{y \to Y(x)^-} \left(f_x(x, y_i) + \int_{y_i}^{y} f_{x,y}(x, \eta)d\eta \right) \\ &= f_x(x, y_i) + \int_{y_i}^{Y(x)} f_{x,y}(x, \eta)d\eta. \end{aligned} \tag{7}$$

Thus for such an x, $f_x(x, Y(x))$ has been defined. It is known from Lemma I that $f_x(x, y_i)$ is measurable and has a summable square and there is a K_i such that

$$\int_{x_i}^{x_{i+1}} | f(x, y_i) |^2 dx \leqq K_i \|f\|^2. \tag{8}$$

Now from (7) $f_x(x, Y(x))$ is a measurable function of x with a summable square, since it is the sum of two functions having this property, and from (6), (7) and (8) using the triangle theorem, we obtain

$$\int_{x_i}^{x_{i+1}} | f_x(x, Y(x)) |^2 dx \leqq K \|f\|^2.$$

Since $Y'(x)$ is monotonically increasing, it follows that if $(x, y) \epsilon$ arc $P_i P_{i+2}$, then there is a relationship $x = X'(y)$ between the coordinates, which is also monotonically increasing and single-valued except possibly for a denumerable set of points, where all values between the limits on the right and left are assumed.

Lemma III. *Let* arc P_iP_{i+2} *be as in Lemma* II. *Let* $X(y)$ *be defined as the greatest number* x, *such that* $Y(x)=y$. $X(y)$ *is a single-valued and monotonically increasing function of* y. *Let* $f_x(X(y), y)$ *be defined as* $\lim_{x\to X(y)^+} f_x(x, y)$ *when the latter exists. Then* $f_x(X(y), y)$ *is defined almost everywhere for the interval* $y_i \leqq y \leqq y_{i+2}$ *and on this interval is measurable and has a summable square and there is a* K *such that*

$$\int_{y_i}^{y_{i+2}} |f_x(X(y), y)|^2 dy \leqq K\|f\|^2.$$

The proof is, by the remark which precedes the lemma, quite analogous to the proof of Lemma II.

Lemma IV. *If* $(x_0, y_0) \epsilon \operatorname{arc} P_iP_{i+2}$ *is such that* $f_x(x_0, Y(x_0))$ *and* $f_x(X(y_0), y_0)$ *both exist, then* $f_x(x_0, Y(x_0)) = f_x(X(y_0), y_0)$.

First suppose that $y_0 \neq y_i$. Now since

$$f_x(x_0, Y(x_0)) = \lim_{y\to Y(x_0)^-} f_x(x_0, y),$$

given an $\epsilon > 0$, we can find a δ' such that for $Y(x_0) - y \geqq Y(x_0) - \delta'$

$$(1) \qquad |f_x(x_0, Y(x_0)) - f_x(x_0, y)| \leqq \epsilon.$$

Similarly we can find a δ'' such that for $X(y_0) < x \leqq X(y_0) + \delta''$,

$$(2) \qquad |f_x'(X(y_0), y_0) - f_x(x, y_0)| \leqq \epsilon.$$

Now for almost every x such that $x_i \leqq x \leqq x_{i+2}$ and all y such that $y_0 > y \geqq y_i$,

$$(3) \qquad \begin{aligned} |f_x(x, y) - f_x(x, y_0)|^2 &= \left|\int_{y_0}^{y} f_{x,y}(x, \xi) d\xi\right|^2 \\ &\leqq |y_0 - y| \left|\int_{y_0}^{y} |f_{x,y}(x, \xi)|^2 d\xi\right|. \end{aligned}$$

Let

$$(4) \qquad g(x, y) = \left|\int_{y_0}^{y} |f_{x,y}(x, \xi)|^2 d\xi\right|.$$

Now $g(x, y) \geqq 0$. Let Δ_y be the rectangle $x_0 \leqq x \leqq x_{i+2}$, $y \leqq \xi < y_0$. Let

$$(5) \qquad \begin{aligned} F(y) &= \int_{x_0}^{x_{i+2}} g(x, y) dx = \int_{x_0}^{x_{i+2}} \left|\int_{y_0}^{y} |f_{x,y}(x, \xi)|^2 d\xi\right| dx \\ &= \iint_{\Delta_y} |f_{x,y}|^2 dS = \left|\int_{y_0}^{y} d\xi \int_{x_0}^{x_{i+2}} |f_{x,y}(x, \xi)|^2 dx\right|. \end{aligned}$$

Then (5) shows that $F(y)\to 0$ as $y\to y_0$. We also have, for y fixed,

$$m_x\{\,|\,y-y_0\,|\,g(x,y)\geqq F(y)^{1/2}\}\leqq|\,y_0-y\,|\,F(y)^{1/2}. \tag{6}$$

Now since $F(y)\to 0$ as $y\to y_0$, we can choose a $\delta'''>0$ in such a way that $F(y)^{1/2}<\epsilon$ for $|y-y_0|<\delta'''$.

Now similarly for almost every y such that $y_0>y\geqq y_i$ and all x such that $x_0<x\leqq x_{i+2}$, we have by the Schwarz inequality,

$$\begin{aligned}|\,f_{x,y}(x,y)-f_x(x_0,y)\,|^2&=\left|\int_{x_0}^{x}f_{x,x}(\eta,y)d\eta\right|^2\\&\leqq|\,x-x_0\,|\int_{x_0}^{x}|\,f_{x,x}(\eta,y)\,|^2d\eta.\end{aligned}\tag{7}$$

Let

$$h=\int_{x_0}^{x}|\,f_{x,x}(\eta,y)\,|^2d\eta,$$

$$\begin{aligned}G(x)=\int_{y_i}^{y_0}dyh(x,y)&=\int_{y_i}^{y_0}dy\int_{x_0}^{x}|\,f_{x,x}(\eta,y)\,|^2d\eta\\&=\iint_{\Delta_x}|\,f_{x,x}(\eta,y)\,|^2dS,\end{aligned}\tag{8}$$

where Δ_x is the rectangle $\{x_0\leqq\eta\leqq x,\ y_i\leqq y\leqq y_0\}$. Now as above $G(x)\to 0$ as $x\to x_0$, and we can choose a δ^{iv} such that for $x_0<x\leqq x_0+\delta^{\text{iv}}$, $G(x)^{1/2}\leqq\epsilon$. Let us suppose that $\epsilon<\frac{1}{4}$. Let $\delta=\min(\delta',\delta'',\delta''',\delta^{\text{iv}})$ and Δ be the rectangle $\{x_0<x\leqq x_0+\delta,\ y_0-\delta\leqq y<y_0\}$. Then from (6) by Fubini's Theorem

$$\begin{aligned}m_{x,y}\{\Delta\{\,|\,y-y_0\,|\,g(x,y)\geqq\epsilon\}\}&\leqq\int_{y_0-\delta}^{y}m_x\{\,|\,y-y_0\,|\,g(x,y)\geqq\epsilon\}dy\\&\leqq\int_{y_0-\delta}^{y_0}|\,y-y_0\,|\,F(y)^{1/2}dy\leqq\int_{y_0-\delta}^{y_0}\delta\epsilon dy\leqq\delta^2\epsilon<\tfrac{1}{4}\delta^2.\end{aligned}\tag{9}$$

Similarly

$$m_{x,y}\{\Delta\cdot\{\,|\,x-x_0\,|\,h(x,y)\geqq\epsilon\}\}<\tfrac{1}{4}\delta^2. \tag{10}$$

Hence the intersection I' of the three sets Δ, the set for which

$$|\,x-x_0\,|\,h(x,y)<\epsilon \tag{11}$$

holds, and the set for which

$$|\,y-y_0\,|\,g(x,y)<\epsilon \tag{12}$$

holds, is by (9) and (10) such that $mI'\geqq\frac{1}{2}\delta^2$.

The intersection I of I' and the sets for which (3) and (7) hold is not empty. Let (x, y) be a point of I. From (3) and (12) we obtain

$$|f_x(x, y) - f_x(x, y_0)| \leqq \epsilon^{1/2}; \tag{13}$$

from (7) and (11),

$$|f_x(x, y) - f_x(x_0, y_0)| \leqq \epsilon^{1/2}. \tag{14}$$

Since $\delta \leqq \min(\delta', \delta'')$, we can apply (1) and (2) which with (13) and (14) yield

$$|f_x(x_0, Y(x_0)) - f_x(X(y_0), y_0)| \leqq 2\epsilon + 2\epsilon^{1/2}.$$

Since ϵ may be taken arbitrarily small we must have

$$f_x(x_0, Y(x_0)) = f_x(X(y_0), y_0).$$

When $y_i = y_0$, we take for the interval for y, when (3) is to be considered, the interval $y_0 > y \geqq y_0 - \eta$, where η is chosen so small that the set (x_0, y) (y in the interval) is in the rectangle associated with P_iP_{i+1}. The proof is then quite similar to the above. When $x_0 = x$ a similar consideration holds.

Definition IV. *$f_x(P)$, $P \in$ arc P_iP_{i+2}, is defined as $f_x(x, Y(x))$ when $P = (x, Y(x))$ and $f_x(P) = f_x(X(y), y)$ when $P = (X(y), y)$.*

From the above lemma we see that this definition is consistent.

Lemma V. *$f_x(P)$ regarded as a function of the arc-length is defined for almost every s which corresponds to a point of arc P_iP_{i+2}, and is measurable for this s interval, and there is a K such that*

$$\int_{\text{arc } P_iP_{i+2}} |f_x(P)|^2 ds \leqq K\|f\|^2.$$

It is well known that for an arc such as P_iP_{i+2}, $x = x(s)$ and $y = y(s)$ are monotonic absolutely continuous functions of s* and that there exist two functions, D_sx and D_sy, defined for every s and such that

$$x = x_i + \int_{s_i}^{s} D_sx\,ds, \quad y = y_i + \int_{s_i}^{s} D_sy\,ds;$$

also such that $D_sx \geqq 0$ and $D_sy \geqq 0$, D_sx being one of the Dini derivatives, and it will be convenient to suppose in the case treated here that it is D_s^+x.

Now

$$s_1 - s_0 = \limsup \sum_1^N (\Delta x_i^2 + \Delta y_i^2)^{1/2},$$

* Cf. Hobson, *Functions of a Real Variable*, pp. 338–341, especially p. 340, also pp. 596, 411.

and since $x(s)$ and $y(s)$ are monotonic,

$$s_1 - s_0 \leqq \limsup \sum_1^N (\Delta x_i + \Delta y_i) = \limsup (x_1 - x_0 + y_1 - y_0) = x_1 - x_0 + y_1 - y_0$$

or

$$1 \leqq \frac{\Delta x}{\Delta s} + \frac{\Delta y}{\Delta s} \cdot$$

Hence we have for every s

$$1 \leqq D_s x + D_s y. \tag{1}$$

Now*

$$\int_{x_i}^{x_{i+2}} |f_x(P)|^2 dx = \int_{\mathrm{arc}\,P_iP_{i+2}} |f_x(P)|^2 D_s x ds, \tag{2}$$

$$\int_{y_i}^{y_{i+2}} |f_x(P)|^2 dy = \int_{\mathrm{arc}\,P_iP_{i+2}} |f_x(P)|^2 D_s y ds, \tag{3}$$

and $|f_x(P)|^2 D_s x$ and $|f_x(P)|^2 D_s y$ are measurable summable functions of s defined for almost every s, corresponding to a point of arc P_iP_{i+2}. (They are defined as zero, when $D_s x$ and $D_s y$ are zero respectively.) Since $D_s x + D_s y$ is a measurable function of s on the s interval corresponding to arc P_iP_{i+2}, $1/(D_s x + D_s y)$ is bounded and measurable on this interval and hence

$$\frac{1}{D_s x + D_s y}\left(|f_x(P)|^2 D_s x + |f_x(P)|^2 D_s y\right) = |f_x(P)|^2$$

is defined almost everywhere and is measurable and summable. Furthermore by (2) and (3) and Lemmas II and III,

$$\begin{aligned}\int_{\mathrm{arc}\,P_iP_{i+2}} |f_x(P)|^2 ds &\leqq \int_{\mathrm{arc}\,P_iP_{i+2}} |f_x(P)|^2 (D_s x + D_s y) ds \\ &\leqq (K_1 + K_2) \|f\|^2.\end{aligned}$$

If the argument of the preceding paragraph is repeated with $f_x(P)$ substituted for $|f_x(P)|^2$, since we have shown that $f_x(x, Y(x))$ and $f_x(X(y), y)$ are measurable functions of x and y respectively, in the appropriate interval, we shall obtain that $f_x(P)$ is a measurable function of s. This completes the proof of the lemma.

Let us suppose now that arc P_jP_{j+1} is an arc of B for which case (2) holds. The proof of the following lemma is quite similar to the proofs of Lemmas II

* Cf. Hobson, loc. cit., p. 665.

and III above and the reader will find no difficulty in making the slight adjustments necessary. We state it for case (2a).

LEMMA VI. *Let* arc P_jP_{j+1} *be such that* (2a) *holds. Then if* $\lim_{x\to X(y)} f_x(x, y)$ *for* $y_j \leqq y \leqq y_{j+1}$, $(x, y)\epsilon S$, *is denoted by* $f_x(X(y), y)$ *when it exists,* $f_x(X(y), y)$ *is a measurable function of* y *and there is a* K *such that*

$$\int_{y_j}^{y_{j+1}} |f_x(X(y), y)|^2 dy \leqq K\|f\|^2.$$

LEMMA VII. *Let* $f_x(P)$, $P\epsilon$ arc P_jP_{j+1}, *be defined as* $f_x(X(y), y)$ *when the latter exists.* $f_x(P)$, $P\epsilon$ arc P_jP_{j+1}, *is a measurable function of* s *defined for almost every* s *corresponding to a point of* arc P_jP_{j+1}, *and there is a* K *such that*

$$\int_{\text{arc}P_jP_{j+1}} |f_x(P)|^2 ds \leqq K\|f\|^2.$$

It is readily seen that $y(s)$ is an absolutely continuous function of s with bounded Dini derivatives and that, conversely, $s(y)$ is such a function of y. Then in view of Lemma VII, (S) Lemma 6.4 (1) will imply by a well known result of the theory of changing the variable in a definite integral*

$$\int_{\text{arc}P_jP_{j+1}} |f_x(P)|^2 ds = \int_{\text{arc}P_jP_{j+1}} D_y s\, |f_x(P)|^2 dy \leqq K\|f\|^2.$$

We have shown that for certain possible situations for cases (1) and (2) of the arcs of B, we may define $f_x(P)$, $P\epsilon B$, for almost every value of s in the appropriate interval, that $f_x(P)$ is measurable with respect to s and $|f_x(P)|^2$ is summable and that its integral along a piece of the arc is less than a constant times $\|f\|^2$. Similar considerations hold for f_y and f and it is quite obvious that the other possibilities which may arise under cases (1) and (2) can be treated in a similar way, and since B is made up of a finite number of such arcs which satisfy (1) or (2) we may conclude

LEMMA VIII. *When* $f_x(P)$, $f_y(P)$ *and* $f(P)$ *are defined as the limits of* f_x, f_y *and* f *respectively along a line parallel to an axis in the manner described above, then* $f_x(P)$, $f_y(P)$, $f(P)$ *are measurable functions of* s *defined for almost every* s *corresponding to a point of* B, *with a summable square, and there exist constants* K_x, K_y *and* K_0 (*independent of* f) *such that*

$$\int_B |f_x(P)|^2 ds \leqq K_x\|f\|^2, \quad \int_B |f_y(P)|^2 ds \leqq K_y\|f\|^2,$$
$$\int_B |f(P)|^2 ds \leqq K_0\|f\|^2.$$

* (S) Lemma 6.3.

10. Boundary value problem

Definition V. *Let the space $\mathfrak{G}$ be defined as the space of classes of functions defined on the boundary and having a summable square with respect to s, two functions f' and g' belonging to the same class if and only if*

$$\int_B |f' - g'|^2 ds = 0.$$

Addition and multiplication by a constant are to have their usual significance, and for $f' \epsilon f$ and $g' \epsilon g$, $(f, g) = \int_B f' g'^{-1} ds$.

That $\mathfrak{G}$ is a Hilbert space is a special case of (S) Theorem 1.24, when we take for E the set, on the s axis, which corresponds to points of B.

Theorem IX. *Let*

$$Q'f' = \beta_1 f_x'(P) + \beta_2 f_y'(P) + \beta_3 f'(P),$$

where the β's are essentially bounded measurable functions of s defined along B. Let Q be the transformation from $\mathfrak{W}$ to $\mathfrak{G}$, which is such that $\{f, g\} \epsilon Q$, if and only if for all functions $f' \epsilon H_s$ and $f' \epsilon f$, $Q'f' \epsilon g$. Then Q is a limited transformation with domain $\mathfrak{W}$.

Since the β's are essentially bounded, measurable functions if $f \epsilon \mathfrak{W}$ is given, it follows from Lemma VII that for every $f' \epsilon f$, $Q'f'$ exists and there is a $g \epsilon \mathfrak{G}$, such that $Q'f' \epsilon g$. Now if $f'' \epsilon f$, then $f' - f'' =_{m+} 0$ and $f' - f''$ is zero for almost every line parallel to an axis. Hence $Q'f' - Q'f'' =_m 0$ and $Q'f'' \epsilon g$. Thus the domain of Q is $\mathfrak{W}$.

Now let M be a measurable upper bound of the absolute values of β_i, $i = 1, 2, 3$. Then

$$\|Qf\|^2 = \int_B |Q'f'|^2 ds \leqq \int_B (|\beta_1|^2 + |\beta_2|^2 + |\beta_3|^2)$$
$$\cdot (|f_x'(P)|^2 + |f_y'(P)|^2 + |f(P)|^2) ds \leqq 3M^2(K_x + K_y + K_0)\|f\|^2,$$

and Q is limited.

We now state our fundamental boundary value for which we can now give a method of solution.

Problem. *Let $L(u)$ be as in the Introduction, $Q'(u)$ as in the preceding Theorem, and let $\overline{\mathfrak{M}}(\mathfrak{R}_0) = \mathfrak{G}'$. Let v be a member of an element of $\mathfrak{K}$, w a member of $\mathfrak{G}'$. Required to find all $u \epsilon H_s$, such that*

$$L(u) =_m v, \quad Q'(u) =_m w. \tag{1}$$

Let E_1' be the projection on the manifold of zeros of the transformation T

associated with L and E_2' that on the manifold of zeros of Q. We can by applying Theorem II construct $I-E_1'=E_1$ and $I-E_2'=E_2$ by means of the method given in the note on construction, since T and Q are limited. On applying Theorem II again, we can state that either no u satisfying (1) exists or u belongs to an element of $\mathfrak{W}$, f, whose projections E_1f and E_2f we can calculate. Furthermore, we know that any u which is in a class f having the calculated projections satisfies (1).

Thus our problem becomes to find all elements f whose projections on two manifolds E_1f and E_2f are given, when we know a complete orthonormal set in each manifold. Or given E_1, E_2, f_1 and f_2, find all f such that

(a) $$E_1f = f_1,$$

(b) $$E_2f = f_2.$$

From (a) we obtain

$$f = f_1 + (I - E_1)f.$$

Substituting in (b) yields

(2) $$E_2(I - E_1)f = f_2 - E_2f_1.$$

Now consider the closed linear manifold $\mathfrak{M}$, determined by the range of $E_2(I-E_1)$. Since $E_2(I-E_1)$ is limited, we can, by using Theorem II, obtain an orthonormal set which determines $\mathfrak{M}$. Hence we can determine whether $f_2-E_2f_1$ is in $\mathfrak{M}$ or not. If $f_2-E_2f_1$ is not in $\mathfrak{M}$ then (2) has no solutions. If $f_2-E_2f_1$ is in $\mathfrak{M}$, since $E_2(I-E_1)$ is limited, taking $\mathfrak{M}$ as our range space, we can apply Theorems 1.16 and II and obtain all f which satisfy (2). Since the manifold of zeros of $E_2(I-E_1)$ includes the range of E_1, we can find all f which satisfy (a) and (2) or the equivalent set of equations (a) and (b).

If, in the above problem, w is zero, we may shorten the above discussion by taking as our domain space for T the closed linear manifold of all f such that $Qf=0$. A similar remark applies, when v is zero.

Columbia University,
New York, N.Y.

THE ASYMPTOTIC FORMS OF THE HERMITE AND WEBER FUNCTIONS*

BY

NATHAN SCHWID

1. **Introduction.** The classical Hermite equation,

$$U''(z) - 2zU'(z) + 2\kappa U(z) = 0, \tag{1}$$

which is satisfied by the Hermite polynomials

$$U_\kappa = (-1)^\kappa e^{z^2} \frac{d^\kappa(e^{-z^2})}{dz^\kappa} \tag{2}$$

when the parameter κ is a positive integer, has been widely discussed. The forms of its solutions, with respect to their asymptotic dependence upon κ, are of importance, and have been determined under certain restrictions upon the variables z and κ. These restrictions, when heaviest, have confined z to real and κ to positive integral values; when lightest, they have permitted z to vary in a strip of the complex plane of finite length and width, and κ over the real axis. In the present paper it is purposed to remove these restrictions: to derive asymptotic forms of the solutions of the equation (1) valid in the entire z plane for large values of κ, real or complex.

It may be recalled that the polynomials (2) were introduced into analysis by Hermite† in 1864. Five years later, Weber‡ noted that the harmonic functions applicable to the parabolic cylinder satisfy an ordinary differential equation of the form

$$w''(z) + (2\kappa + 1 - z^2)w(z) = 0, \tag{3}$$

which has since been generally known as the Weber equation. Whittaker§ showed in 1903 that this equation is obtainable from the Hermite equation (1) by a simple change of variable, and he determined the asymptotic expansion with respect to the *real* variable z, of a particular solution, the classic $D_\kappa(z)$.

* Presented to the Society, April 6, 1934; received by the editors May 17, 1934.

† Hermite, *Sur un nouveau développement en série des fonctions*, Comptes Rendus, vol. 58, pp. 93–100.

‡ Weber, *Ueber die Integration der partiellen Differentialgleichung*: $\partial^2u/\partial x^2+\partial^2u/\partial y^2+k^2u=0$, Mathematische Annalen, vol. 1 (1869), pp. 1–36.

§ Whittaker, *On the functions associated with the parabolic cylinder in harmonic analysis*, Proceedings of the London Mathematical Society, vol. 35, pp. 417–427.

Soon thereafter, Adamoff* obtained the asymptotic forms of the *Hermite polynomials* relative to the integral parameter κ, with the variable z limited to *real*, *finite* values.

Watson† generalized these results. He developed the asymptotic expansion of $D_\kappa(z)$ with respect to z for *all* values of arg z, and also determined, in a strip of the z plane of finite width and length, the asymptotic forms of D_κ relative to the *real* parameter κ. His method, in the latter case, was a generalization of Adamoff's, whose procedure was based upon an elaborate transformation of a definite integral. More recently, Plancherel and Rotach‡ derived asymptotic forms for the *Hermite polynomials* with respect to the integral parameter κ, for *all real values* of z. They used the saddle-point method§ applied to a contour integral

$$H_{\kappa-1}(x) = \frac{(\kappa-1)!}{(-1)^{\kappa-1}} \int_C \frac{e^{-z^2/2-zx}}{z^\kappa} dz.$$

This procedure, though a familiar one, is not so intimately connected with the differential equation.‖

In the present paper, the asymptotic forms of the solutions of the Hermite and Weber equations with respect to the *complex* parameter κ are obtained for *all* complex values of z. This is done by utilizing formulas developed by Langer¶ for the asymptotic solutions of an ordinary differential equation of the general structure

$$w''(z) + p(z)w'(z) + \{\rho^2\phi^2(z) + q(z)\} w(z) = 0,$$

where the parameter ρ^2 is large and the variable z ranges over some region (finite or infinite) of the z plane, in which the coefficient ϕ^2 vanishes to some real non-negative power at one and only one point.

* Adamoff, *Sur les expansions des polynomes $U_n = e^{\alpha x^2/2} d^n e^{-\alpha x^2/2}/dx^n$ pour les grandes valeurs de n*, Annales de l'Institut Polytechnique de St. Petersburg, 1906, pp. 127–143.

† Watson, G. N., *The harmonic functions associated with the parabolic cylinder*, Proceedings of the London Mathematical Society, (2), vol. 8, pp. 393–421.

‡ Plancherel et Rotach, *Sur les valeurs asymptotiques des polynomes d'Hermite*, Commentarii Mathematici Helvetici, vol. 1 (1929), pp. 227–254.

§ For a discussion of this method, see Courant-Hilbert, *Methoden der Mathematischen Physik*, vol. I, p. 435.

‖ Some additional related material of interest can be found in the following articles:

E. Hille, *On the zeros of the functions of the parabolic cylinder*, Arkiv för Matematik, Astronomi och Fysik, vol. 18 (1924), No. 26.

R. Nevanlinna, *Über Riemannsche Flächen mit endlich vielen Windungspunkten*, Acta Mathematica, vol. 58 (1932), pp. 295–373, especially pp. 344–355 and 361–372.

¶ Langer, R. E., *On the asymptotic solutions of differential equations*, etc., these Transactions, vol. 34, No. 3, pp. 447–480.

The asymptotic formulas here derived, which include as special cases the above mentioned forms of Watson (relative to large κ) and of Plancherel and Rotach, are shown to be in accord with the work of these men.

2. **Preliminary considerations.** The change of variable

$$U = we^{z^2/2} \tag{4}$$

relates the equations (1) and (3), and in the latter, the substitutions

$$z = (2\kappa + 1)^{1/2}(t + 1), \ \rho = i(2\kappa + 1), \ w(z) \equiv u(t) \tag{5}$$

further reduce the equation to the form

$$u''(t) + \rho^2(t^2 + 2t)u(t) = 0. \tag{6}$$

This equation is of the type

$$u'' + \rho^2\phi^2 u = 0,$$

for the solutions of which asymptotic formulas have been found by Langer.*

The equation (3) is unchanged when z is replaced by $(-z)$, so that its principal solutions at the origin, which will be designated $w_1(z)$ and $w_2(z)$, are, respectively, even and odd functions. The variable z may, accordingly, be restricted to some half plane, a convenient choice being

$$\begin{gathered} -\frac{\pi}{2} + \frac{1}{2}\arg(2\kappa + 1) < \arg z \leqq \frac{\pi}{2} + \frac{1}{2}\arg(2\kappa + 1), \text{ that is,} \\ R(t) \geqq -1. \end{gathered} \tag{7}$$

This part of the t plane, cut along the negative real axis, so that

$$-\pi < \arg t \leqq \pi, \tag{8}$$

will be referred to as R_t, and the corresponding region (7) in the z plane, cut from the origin to the point $z=(2\kappa+1)^{1/2}$, will be referred to as R_z.

The function $\phi=(t^2+2t)^{1/2}$ is evidently single-valued in R_t, and is completely specified if ϕ is chosen so that $\arg\phi=0$ when $\arg t=0$.

The complex parameter κ is to be thought of as unbounded in magnitude but bounded from zero; its argument is restricted to the range of values

$$-\frac{3\pi}{2} < \arg(2\kappa + 1) \leqq \frac{\pi}{2}, \ \text{that is,} \ -\pi < \arg\rho \leqq \pi. \tag{9}$$

The asymptotic forms given in [L]† were obtained under certain hy-

* Langer considered the more general equation, $u''+(\rho^2\phi^2-\chi)u=0$, in which χ is analytic. In equation (6), $\chi\equiv 0$.

† This abbreviation will be used hereafter in referring to the previously mentioned paper of Langer.

potheses on the coefficients of the differential equation, and a brief examination of these will suffice to show that they are satisfied by the equation (6).

(i) It is required that ϕ^2 be of the structure $t^\nu\phi_1$, where $\nu \geqq 0$, and ϕ_1 is single-valued, analytic, and non-vanishing in R_t; here this is obviously fulfilled, with $\phi^2 = t(t+2)$.

(ii) The function $\Phi \equiv \int_0^t \phi(t)dt$, which, in this case, is of the specific form

$$\Phi = \tfrac{1}{2}\{(t+1)(t^2+2t)^{1/2} - \log\,[t + 1 + (t^2+2t)^{1/2}]\}, \tag{10}$$

is to be non-vanishing in R_t except at the origin.

That Φ as given in (10) fulfills this requirement is readily shown. In terms of the complex variable θ defined by the relation $t + 1 = \cos\theta$, the formula (10) can be written

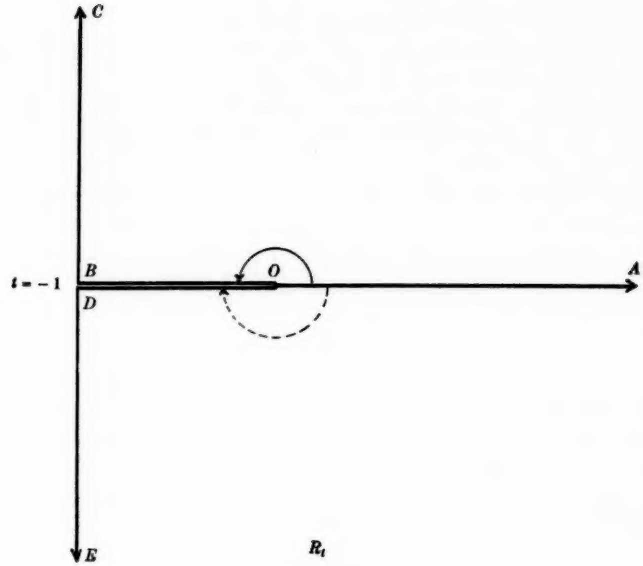
$$\Phi = \pm \frac{i}{4}(\sin 2\theta - 2\theta). \tag{11}$$

From this relation it is easily verified that Φ vanishes only when $\theta = 0$, i.e., when $t = 0$.

Before examining the remaining hypotheses, it is desirable to discuss the map of the region R_t upon the Φ plane. From the definition of Φ, it is readily seen that the map is conformal, except at the origin, where the ratio of angles in Φ and t planes is as 3:2. The manner in which the formula (10) determines the map is made apparent from the following table I; R_t and its map R_Φ are shown in Figures 1 and 2 respectively.

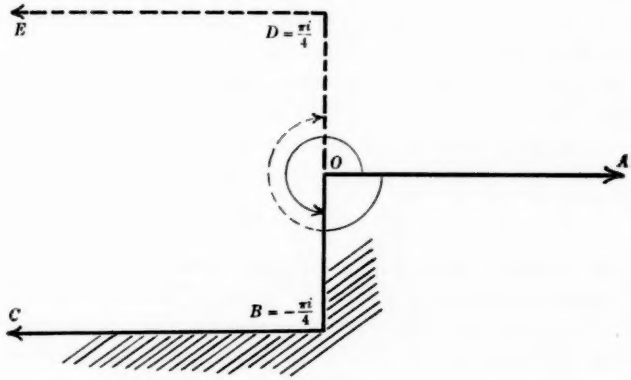
TABLE I

Line	t	Φ
OA	0 to ∞	0 to ∞
OB	0 to -1	0 to $-\dfrac{\pi i}{4}$
BC	-1 to $-1 + i\infty$	$-\dfrac{\pi i}{4}$ to $-\infty - \dfrac{\pi i}{4}$
OD	0 to -1	0 to $\dfrac{\pi i}{4}$
DE	-1 to $-1 - i\infty$	$\dfrac{\pi i}{4}$ to $-\infty + \dfrac{\pi i}{4}$



R_t

FIGURE 1



R_ϕ

FIGURE 2

The consideration of the hypotheses may now be continued.

(iii) This, the hypothesis (iv) of [L], is a condition upon R_ξ, the map of R_t upon the ξ plane, where $\xi = \rho\Phi$. Without a consideration of details, it may be stated that, roughly, this hypothesis requires that the surface R_ξ be one that contains in its interior only one segment of any horizontal line $\mathfrak{J}(\xi) = $ a constant. Since R_ξ is obtained from R_Φ by a rotation about the origin and a magnification of the unit in the ratio $|2\kappa+1|:1$, it is evident from Figure 2 that the surface R_ξ is such that the hypothesis is satisfied.

(iv) The region R_t is to be such that for some positive number M, the relation

$$\int \left| \frac{\Psi''}{\Psi\phi} dt \right| < M, \tag{12}$$

where $\Psi = \Phi^{1/6}/\phi^{1/2}$, is uniformly valid with respect to integrations over all arcs upon which $\mathfrak{J}(\xi)$ varies monotonically with $|\xi|$, and upon which $|t| \geqq N$, for some positive number N.

From the definition of Φ and the formula (10), it follows that Φ is of the structure

$$\begin{aligned} \Phi &= \frac{(2t)^{3/2}}{3}[1 + o(t)], \quad |t| < 2, \\ \Phi &= \frac{t^2}{2}[1 + O(t^{-1})], \quad |t| > 2, \end{aligned} \tag{13}$$

and by using the second of these relations, it may be readily verified that the integrand of (12) is $O(t^{-3})$ when $|t| > 2$. Hence, if N is sufficiently large, the hypothesis is satisfied for all arcs of the required type.

The equation (6) is thus one to which the theorems proved in [L] are applicable.

3. **The regions of validity.** Let the regions $\Xi^{(h)}$ be defined by the relation

$$(h-1)\pi + \epsilon \leqq \arg \xi \leqq (h+1)\pi - \epsilon \quad (h = 0, \pm 1, \pm 2, \cdots), \tag{14}$$

where ϵ is an arbitrary positive fixed constant sufficiently small. Asymptotic formulas, dependent upon h, are given in [L] for a certain set of independent solutions of the equation (6), and for any given h, the associated set of formulas is, by [Theorem 7, L], valid in $\Xi^{(h)}$; i.e., valid for all values of t for which $\xi(t)$ satisfies the relations (14). The particular values of the indices which will fix the regions for the present problem may be determined by noting that in R_Φ, Φ is restricted to the values

$$-\frac{3\pi}{2} < \arg \Phi \leqq \frac{3\pi}{2},$$

so that in R_ξ,

$$\arg \rho - \frac{3\pi}{2} < \arg \xi \leqq \arg \rho + \frac{3\pi}{2}\cdot$$

When κ and its counterpart ρ vary respectively in the half planes

$$-\frac{\pi}{2} < \arg (2\kappa + 1) \leqq \frac{\pi}{2},$$

$$0 < \arg \rho \leqq \pi,$$

ξ is restricted to the values

$$-\frac{3\pi}{2} < \arg \xi \leqq \frac{5\pi}{2};$$

this range comprises, wholly or in part, the regions $\Xi^{(h)}$, $h = -1, 0, 1, 2$, with boundaries which may be conveniently specified in terms of Φ as follows:

$0 < \arg \rho \leqq \pi$:

$$\begin{aligned}
&\Xi^{(2)}: && \pi - \arg \rho + \epsilon \leqq \arg \Phi \leqq \frac{3\pi}{2},\\
&\Xi^{(1)}: && -\arg \rho + \epsilon \leqq \arg \Phi \leqq 2\pi - \arg \rho - \epsilon,\\
(15)\qquad &\Xi^{(0)}: && -\pi - \arg \rho + \epsilon \leqq \arg \Phi \leqq \pi - \arg \rho - \epsilon,\\
&\Xi^{(-1)}: && -\frac{3\pi}{2} \leqq \arg \Phi \leqq -\arg \rho - \epsilon.
\end{aligned}$$

Similarly, when κ and ρ are restricted to the respective half planes

$$-\frac{3\pi}{2} < \arg (2\kappa + 1) \leqq -\frac{\pi}{2},$$

$$-\pi < \arg \rho \leqq 0,$$

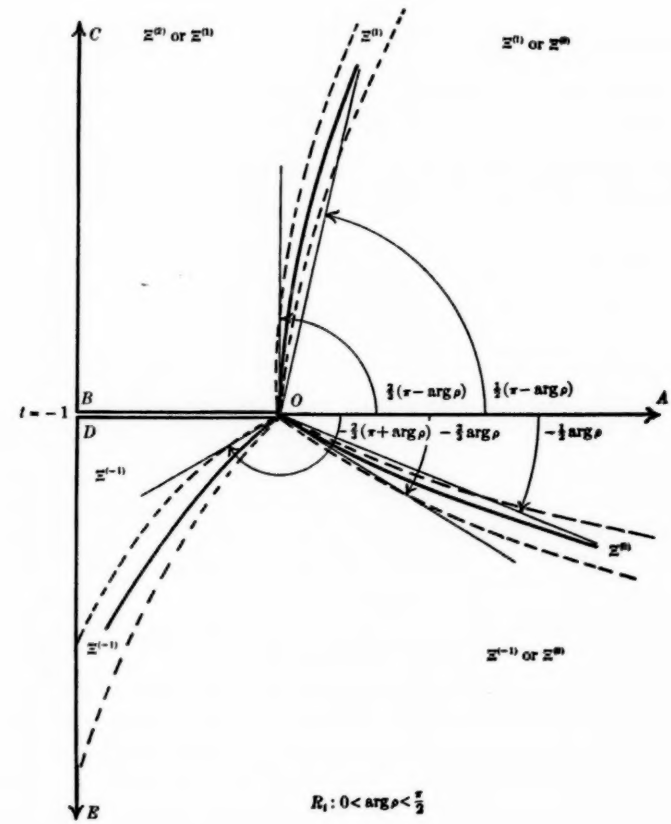
ξ is confined by the bounds

$$-\frac{5\pi}{2} < \arg \xi \leqq \frac{3\pi}{2},$$

for which the corresponding regions are those with the indices $-2, -1, 0, 1$; these have the boundaries

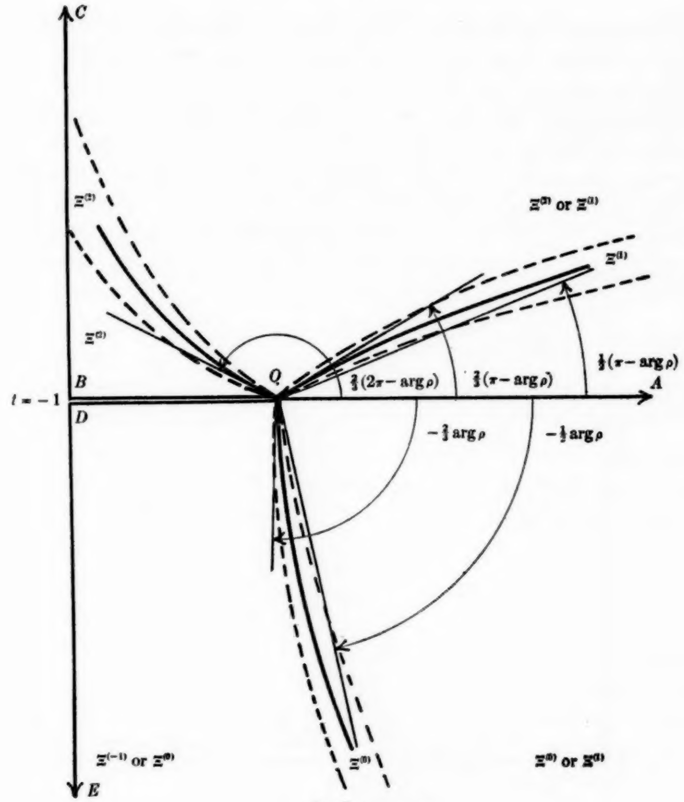
$-\pi < \arg \rho \leqq 0$:

$$(16)\qquad \begin{aligned}
&\Xi^{(-2)}: && -\frac{3\pi}{2} < \arg \Phi \leqq -\pi - \arg \rho - \epsilon,\\
&\Xi^{(-1)}: && -2\pi - \arg \rho + \epsilon \leqq \arg \Phi \leqq -\arg \rho - \epsilon,\\
&\Xi^{(0)}\ : && -\pi - \arg \rho + \epsilon \leqq \arg \Phi \leqq \pi - \arg \rho - \epsilon,\\
&\Xi^{(1)}\ : && -\arg \rho + \epsilon \leqq \arg \Phi \leqq \frac{3\pi}{2}\cdot
\end{aligned}$$



$R_1: 0 < \arg \rho < \frac{\pi}{2}$

FIGURE 3(a)



$R_t : \frac{\pi}{2} < \arg\rho \leqq \pi$

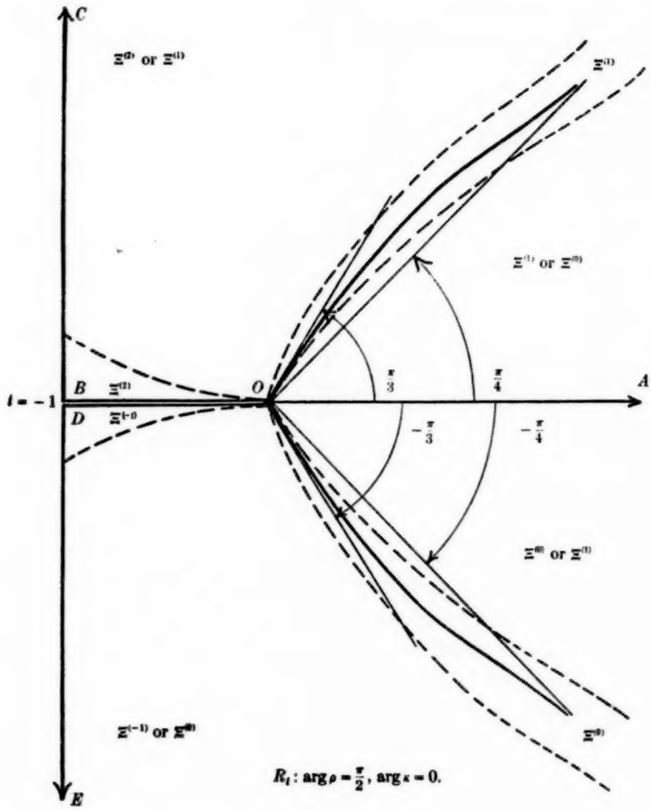
FIGURE 3(b)

The manner in which the boundary lines of (15) and (16) map as curves in R_t becomes evident on examining the formulas (13); for from these it is seen that $\arg t \to \frac{1}{2} \arg \Phi$ when $t \to \infty$, and $\arg t \to \frac{2}{3} \arg \Phi$ when $t \to 0$. Hence a radial line $\arg \Phi = \alpha$ maps as a curve in R_t tangent at $t=0$ to the line $\arg t = 2\alpha/3$, and asymptotic to the radial line $\arg t = \alpha/2$.

The regions $\Xi^{(h)}$ when $\arg \rho > 0$ are shown in Figure 3 for the typical cases $0 < \arg \rho < \pi/2$, $\pi/2 < \arg \rho \leqq \pi$, and the important special case, $\arg \rho = \pi/2$, $\arg \kappa = 0$. To obtain the corresponding configurations when $\arg \rho < 0$, it is only necessary to reflect Figure 3 in the axis of reals and to change the signs of the indices of the regions; these latter configurations,

although not specifically given here, will, for convenience, be referred to as Figure 4. The boundary curves in R_s are readily obtained from Figures 3 and 4 by the appropriate change of unit, rotation and translation.

It will be noted that the regions overlap for consecutive values of h. A set of asymptotic formulas will be given for each region, and in the part common to two, either of the associated sets of formulas may be used, since they are asymptotically equivalent. Hence the exact location of the boundary curves of the regions (the dotted lines of Figure 3) is not important, for if at any point the validity of one form is in doubt, it will be certain that the other of the two formulas is then valid.



R_t : $\arg \rho = \frac{\pi}{2}$, $\arg \kappa = 0$.

FIGURE 3(c)

The forms derived in §4 are subject to the restriction that $|\rho|$ and $|\xi|$ be large; subsequently, forms are derived for bounded values of $|\xi|$, $|\rho|$ large. From the structure of Φ as shown in equations (13), it is evident that $|\xi|$ is bounded when $t=O(|\rho|^{-2/3})$, or, in terms of z, when $z=O(|\rho|^{1/2})$. Hence the asymptotic forms for large ξ are valid in the regions $\Xi^{(h)}$ as shown in Figures 3 and 4 with the exception of a region of dimension $O(|\rho|^{-2/3})$ about the point $t=0$; i.e., about $z=(2\kappa+1)^{1/2}$; in this latter region, the forms obtained for bounded values of ξ apply.

Figure 4(a): R_t, $-\pi/2<\arg \rho\leqq 0$; this is Figure 3(a) reflected in real axis, and signs of indices changed.

Figure 4(b): R_t, $-\pi<\arg \rho<-\pi/2$; Figure 3(b) reflected in real axis, signs of indices changed.

Figure 4(c): R_t, $\arg \rho=-\pi/2$, $\arg \kappa=-\pi$; Figure 3(c) reflected in real axis, signs of indices changed.

4. **The asymptotic forms of the Weber functions.** The solutions $w_j(z)$ of the equation (3) are expressible linearly in terms of any independent set of solutions, $u_j(t)$, of the equation (6), and these expressions, identities in t, may be written

$$w_j(z) \equiv A_{1j}^{(k)} u_1(t) + A_{2j}^{(k)} u_2(t) \qquad (j = 1, 2), \tag{17}$$

in which the superscript k has a significance which will presently be explained. The quantities $A_{ij}^{(k)}$ are constants with respect to t, but may depend upon ρ, so that any value of t in R_t may be used for their determination. The identities (17) reduce, when $t=-1$, to linear systems in $A_{ij}^{(k)}$:

$$\begin{aligned} &A_{11}^{(k)} u_1(-1) + A_{21}^{(k)} u_2(-1) = 1,\\ &A_{11}^{(k)} u_1'(-1) + A_{21}^{(k)} u_2'(-1) = 0; \end{aligned} \tag{18a}$$

$$\begin{aligned} &A_{12}^{(k)} u_1(-1) + A_{22}^{(k)} u_2(-1) = 0,\\ &A_{12}^{(k)} u_1'(-1) + A_{22}^{(k)} u_2'(-1) = (2\kappa + 1)^{1/2}. \end{aligned} \tag{18b}$$

To solve these systems, the quantities $u_j(-1)$ and $u_j'(-1)$ must be computed. Asymptotic forms are given in [L] for an independent set of solutions of equation (6), which, together with their derivatives, assume at $t=0$ certain values.* These values are not, in themselves, germane to the present problem and are not here reproduced, but the functions u_j of (17) are to be thought of as those solutions of (6) which take on at $t=0$ values as prescribed

* Cf. [L], p. 460, equations (23).

in [L] and for which asymptotic formulas are there given.* Because of the presence of the Stokes' phenomenon, the asymptotic representation of a solution u_j is different in distinct regions $\Xi^{(h)}$. Hence, in order to compute $u_j(-1)$ and $u_j'(-1)$ it is necessary to know in what region the variable ξ is found when $t=-1$. Since

$$\xi(-1) = \frac{\pi\rho}{4}\exp\left(\frac{3\pi i}{2}\right),$$

it follows that

$$\xi(-1) \text{ is in } \Xi^{(2)},\ 0 < \arg\rho \leqq \pi,$$
$$\xi(-1) \text{ is in } \Xi^{(1)},\ -\pi < \arg\rho \leqq 0,$$

and the superscript k of (17) will be assigned the values 2 or 1 according as ρ is in its upper or lower half plane.

The quantities $u_j(-1)$ and $u_j'(-1)$ are now computed without difficulty from the formulas of [L], their values being†‡

$0 < \arg\rho \leqq \pi$:

$$(19a)\quad \begin{aligned} u_j(-1) &= e^{(\pm 1-1)\pi i/3}\left(\frac{2}{\pi}\right)^{1/2}(2\kappa+1)^{-1/6}\left[\cos\frac{\pi}{2}\left(\kappa \pm \frac{1}{3}\right)\right],\\ u_j'(-1) &= e^{(\pm 1-1)\pi i/3}\left(\frac{2}{\pi}\right)^{1/2}(2\kappa+1)^{5/6}\left[\sin\frac{\pi}{2}\left(\kappa \pm \frac{1}{3}\right)\right], j = 1, 2; \end{aligned}$$

$-\pi < \arg\rho \leqq 0$:

$$(19b)\quad \begin{aligned} u_j(-1) &= e^{(7\mp 2)\pi i/6}\left(\frac{2}{\pi}\right)^{1/2}(2\kappa+1)^{-1/6}\left[\sin\frac{\pi}{2}\left(\kappa \mp \frac{1}{3}\right)\right],\\ u_j'(-1) &= e^{(1\mp 2)\pi i/6}\left(\frac{2}{\pi}\right)^{1/2}(2\kappa+1)^{5/6}\left[\cos\frac{\pi}{2}\left(\kappa \mp \frac{1}{3}\right)\right],\quad j = 1, 2. \end{aligned}$$

From these formulas, the determinant of the systems (18) is computed to be $(2\kappa+1)^{2/3}e^{\pi i/3}3^{1/2}[1]/\pi$, and the systems are found to have the solutions

$$(20)\quad \begin{aligned} A_{11}^{(2)} &= e^{(1\mp 1)\pi i/3}\left(\frac{2\pi}{3}\right)^{1/2}(2\kappa+1)^{1/6}\left[\cos\frac{\pi}{2}\left(\kappa \pm \frac{2}{3}\right)\right],\\ A_{12}^{(2)} &= e^{(1\mp 1)\pi i/3}\left(\frac{2\pi}{3}\right)^{1/2}(2\kappa+1)^{-1/3}\left[\sin\frac{\pi}{2}\left(\kappa \pm \frac{2}{3}\right)\right], \end{aligned}$$

* Ibid., p. 462, formulas (29).

† Double signs used in connection with two indices indicate that the upper sign is to be used with the first, the lower sign with the second index.

‡ The symbol $[Q]$ will always denote a series

$$Q + \sum_{m=1}^{\infty} E_m/\rho^m$$

in which the E_m are bounded functions of ρ and t, or if t is not present, of ρ alone.

$$A_{l1}^{(1)} = e^{(\pm 2-7)\pi i/6}\left(\frac{2\pi}{3}\right)^{1/2}(2\kappa+1)^{1/6}\left[\sin\frac{\pi}{2}\left(\kappa\mp\frac{2}{3}\right)\right],$$

$$A_{l2}^{(1)} = e^{(\pm 2-1)\pi i/6}\left(\frac{2\pi}{3}\right)^{1/2}(2\kappa+1)^{-1/3}\left[\cos\frac{\pi}{2}\left(\kappa\mp\frac{2}{3}\right)\right], \qquad l = 1, 2.$$

The asymptotic forms of $u_j(t)$ when $|\xi|>N$ are given in [L] as

$$u_j(t) \sim \rho^{-1/6}\phi^{-1/2}\{a_{j1}^{(h)}e^{i\xi}S_{1h} + a_{j2}^{(h)}e^{-i\xi}S_{2h}\}, \qquad j = 1, 2,$$

in which the quantities S_{jh} are functions having the general structure $1+O(\rho^{-1})+O(\xi^{-1})$, the quantities a_{ji} are constants which can be computed from formulas in [L, pp. 460–61], and the superscript h denotes the region $\Xi^{(h)}$ in which the formula is valid. The substitution of this and the formulas (20) in the identity (17) leads to the asymptotic forms of $w_j(z)$, and the results obtained are summarized in the following theorems:

THEOREM I. *The asymptotic forms of the principal solutions $w_j(z)$ of the equation* (3), *valid, when $-\pi/2<\arg(2\kappa+1)\leqq\pi/2$, in the regions $\Xi^{(h)}$ as shown in Figure* 3, *are given by the formula*

$$(21)\quad w_j(z) \sim \frac{1}{2(z^2/(2\kappa+1)-1)^{1/4}}\{B_{j1}^{2h}e^{i\xi} + B_{j2}^{2h}e^{-i\xi}\}\ (j = 1, 2; h = 2, 1, 0, -1),$$

in which the functions $\exp(\pm i\xi)$ *are of the specific form*

$$(22)\quad e^{\pm i\xi} = \exp\left(\mp(z^2-2\kappa-1)^{1/2}z/2\right)\left\{\frac{z}{(2\kappa+1)^{1/2}} + \left(\frac{z^2}{2\kappa+1}-1\right)^{1/2}\right\}^{\pm(2\kappa+1)/2}$$

and the quantities B_{jl}^{2h} are functions of κ whose values in the regions $\Xi^{(h)}$ are shown in the appended table II.

TABLE II

h	B_{11}^{2h}	B_{12}^{2h}	B_{21}^{2h}	B_{22}^{2h}
2	$(-1)^{-\kappa/2}[1]$	$(-1)^{(\kappa+1)/2}[1]$	$\frac{(-1)^{(1-\kappa)/2}}{(2\kappa+1)^{1/2}}[1]$	$\frac{(-1)^{\kappa/2}}{(2\kappa+1)^{1/2}}[1]$
1	$(-1)^{-\kappa/2}[1]$	$-2\left[\sin\frac{\pi\kappa}{2}\right]$	$\frac{(-1)^{(1-\kappa)/2}}{(2\kappa+1)^{1/2}}[1]$	$\frac{2}{(2\kappa+1)^{1/2}}\left[\cos\frac{\pi\kappa}{2}\right]$
0	$(-1)^{\kappa/2}[1]$	$-2\left[\sin\frac{\pi\kappa}{2}\right]$	$\frac{(-1)^{(\kappa-1)/2}}{(2\kappa+1)^{1/2}}[1]$	$\frac{2}{(2\kappa+1)^{1/2}}\left[\cos\frac{\pi\kappa}{2}\right]$
-1	$(-1)^{\kappa/2}[1]$	$(-1)^{-(\kappa+1)/2}[1]$	$\frac{(-1)^{(\kappa-1)/2}}{(2\kappa+1)^{1/2}}[1]$	$\frac{(-1)^{\kappa/2}}{(2\kappa+1)^{1/2}}[1]$

THEOREM II. *The asymptotic forms of the principal solutions $w_j(z)$ of the equation* (3) *valid, when* $-3\pi/2<\arg(2\kappa+1)\leqq-\pi/2$, *in the regions* $\Xi^{(h)}$ *as explained in Figure* 4, *are given by the formula*

$$(23)\quad w_j(z)\sim\frac{1}{2\left(\dfrac{z^2}{2\kappa+1}-1\right)^{1/4}}\{B_{j1}^{1h}e^{i\xi}+B_{j2}^{1h}e^{-i\xi}\}$$

$$(j=1,2;\ h=1,0,-1,-2),$$

in which the functions $\exp(\pm i\xi)$ *are described by the formula* (22), *and the quantities* B_{jl}^{1h} *have values in* $\Xi^{(h)}$ *as shown in the table* III.

TABLE III

h	B_{11}^{1h}	B_{12}^{1h}	B_{21}^{1h}	B_{22}^{1h}
1	$(-1)^{-\kappa/2}[1]$	$(-1)^{(\kappa+1)/2}[1]$	$\dfrac{(-1)^{(1-\kappa)/2}}{(2\kappa+1)^{1/2}}[1]$	$\dfrac{(-1)^{\kappa/2}}{(2\kappa+1)^{1/2}}[1]$
0	$2\left[\cos\dfrac{\pi\kappa}{2}\right]$	$(-1)^{(\kappa+1)/2}[1]$	$\dfrac{2\left[\sin\dfrac{\pi\kappa}{2}\right]}{(2\kappa+1)^{1/2}}$	$\dfrac{(-1)^{\kappa/2}}{(2\kappa+1)^{1/2}}[1]$
-1	$2\left[\cos\dfrac{\pi\kappa}{2}\right]$	$(-1)^{-(\kappa+1)/2}[1]$	$\dfrac{2\left[\sin\dfrac{\pi\kappa}{2}\right]}{(2\kappa+1)^{1/2}}$	$\dfrac{(-1)^{-\kappa/2}}{(2\kappa+1)^{1/2}}[1]$
-2	$(-1)^{\kappa/2}[1]$	$(-1)^{-(\kappa+1)/2}[1]$	$\dfrac{(-1)^{(\kappa-1)/2}[1]}{(2\kappa+1)^{1/2}}$	$\dfrac{(-1)^{-\kappa/2}[1]}{(2\kappa+1)^{1/2}}$

5. **Special forms when κ is real.** When the parameter κ is real, fairly simple forms of $w_j(z)$ are valid inside the semi-circle $|z|=|2\kappa+1|^{1/2}$ contained in the half plane (7). When κ is positive, (7) is the half plane

$$R(z)\geqq 0,$$

and the semi-circle is entirely within the regions $\Xi^{(h)}$, $h=2,-1$. Then, by Theorem I, the forms that apply are

$$w_1(z)\sim\frac{1}{2\left(1-\dfrac{z^2}{2\kappa+1}\right)^{1/4}}\left\{\exp\left(i\left(\xi\mp\frac{\pi\kappa}{2}\mp\frac{\pi}{4}\right)\right)[1]\right.$$

$$\left.+\exp\left(-i\left(\xi\mp\frac{\pi\kappa}{2}\mp\frac{\pi}{4}\right)\right)[1]\right\},$$

$$(24) \qquad w_2(z) \sim \frac{1}{2(2\kappa+1)^{1/2}\left(1-\dfrac{z^2}{2\kappa+1}\right)^{1/4}}\left\{\exp\left(i\left(\xi \mp \frac{\pi\kappa}{2} \pm \frac{\pi}{4}\right)\right)[1]\right.$$
$$\left. + \exp\left(-i\left(\xi \mp \frac{\pi\kappa}{2} \pm \frac{\pi}{4}\right)\right)[1]\right\},$$

in which

$$\arg\left(1-\frac{z^2}{2\kappa+1}\right) \leqq 0 \text{ when } y \geqq 0, \ \arg\left(1-\frac{z^2}{2\kappa+1}\right) > 0, \ y < 0.$$

Attention has already been called (equation (11)) to the fact that with the substitution

$$\cos\theta = \frac{z}{(2\kappa+1)^{1/2}},$$

ξ takes the form

$$(25) \qquad \xi = \pm\frac{2\kappa+1}{2}(\theta - \sin\theta\cos\theta),$$

and if the sign of arg $(1-z^2/(2\kappa+1))$ be determined as in (24), then here the upper sign applies when $y \geqq 0$ inside the semi-circle, the lower when $y<0$.

The variable θ, inside the semi-circle, is of the structure

$$\theta = \frac{\pi}{2} - \frac{z}{(2\kappa+1)^{1/2}} + O(\kappa^{-3/2}),$$

so that by (25), ξ may be written

$$(26) \qquad \xi = \pm\left\{(2\kappa+1)\frac{\pi}{4} - z(2\kappa+1)^{1/2} + O(\kappa^{-1/2})\right\}.$$

The use of this formula in (24) gives to the latter the forms

$$w_1(z) \sim \frac{1}{\left(1-\dfrac{z^2}{2\kappa+1}\right)^{1/4}}\{\cos z(2\kappa+1)^{1/2} + O(\kappa^{-1/2})\},$$

$$w_2(z) \sim \frac{1}{(2\kappa+1)^{1/2}\left(1-\dfrac{z^2}{2\kappa+1}\right)}\{\sin z(2\kappa+1)^{1/2} + O(\kappa^{-1/2})\},$$
$$0 \leqq |y| \leqq \epsilon,$$

ϵ being an arbitrary fixed positive quantity sufficiently small.

When $y>\epsilon$, the first of forms (24) may be written, with the aid of (26), in the form

$$w_1(z) \sim \frac{\cos z(2\kappa+1)^{1/2}}{\left(1-\dfrac{z^2}{2\kappa+1}\right)^{1/4}}\left\{1+\frac{E_1+E_2\exp\left(2iz(2\kappa+1)^{1/2}\right)}{1+\exp\left(2iz(2\kappa+1)^{1/2}\right)}\kappa^{-1/2}\right\},$$

in which E_j are bounded functions of z and κ. It follows that the numerator of the fraction inside the brace is also bounded in z and κ. The denominator of this fraction vanishes only on the real axis, at the points

$$x = \frac{2n+1}{(2\kappa+1)^{1/2}}\,\frac{\pi}{2} \qquad (n = 0, 1, 2, \cdots)$$

so that the denominator is non-vanishing if $y>\epsilon$; but more than this, it is then bounded from zero, uniformly in z and κ; for if $2\kappa+1>M^2$, then

$$\begin{aligned} \left|1+\exp\left(2iz(2\kappa+1)^{1/2}\right)\right| &= \left|1-\exp\left(-2y(2\kappa+1)^{1/2}\right)\left(\exp\left(-2y(2\kappa+1)^{1/2}\right)\right.\right.\\ &\qquad \left.\left. +\, 2\cos x(2\kappa+1)^{1/2}\right)\right|, \\ \left|1+\exp\left(2iz(2\kappa+1)^{1/2}\right)\right| &\geqq 1-\exp\left(-2y(2\kappa+1)^{1/2}\right) > 1-e^{-2\epsilon M}>0; \end{aligned}$$

hence $w_1(z)$ may be written, when $y>\epsilon$,

$$w_1(z) \sim \frac{\cos z(2\kappa+1)^{1/2}}{\left(1-\dfrac{z^2}{2\kappa+1}\right)^{1/4}}\left\{1+O(\kappa^{-1/2})\right\}.$$

The same conclusion is reached when $|y|>\epsilon$, and similar results hold for $w_2(z)$.

When κ is real and negative, the half plane (7) is described by the relation

$$\mathfrak{I}(z) \leqq 0,$$

and in the regions $\Xi^{(h)}$, $h=1, -2$, the forms that apply are those given in (24). Inside the semi-circle under consideration, the results that obtain are similar to those obtained when $\kappa>0$. The facts developed in this section are summed up in the following theorems:

THEOREM III. *When the parameter κ is positive, the asymptotic forms of $w_j(z)$ inside the semi-circle $|z|=(2\kappa+1)^{1/2}$ contained in the half plane $R(z)\geqq 0$ are given by the formulas*

$$w_1(z) \sim \frac{1}{\left(1-\dfrac{z^2}{2\kappa+1}\right)^{1/4}}\begin{cases}\cos z(2\kappa+1)^{1/2}+O(\kappa^{-1/2}), & 0\leqq |y| \leqq \epsilon,\\ \cos z(2\kappa+1)^{1/2}(1+O(\kappa^{-1/2})), & |y|>\epsilon,\end{cases}$$

(27)
$$w_2(z) \sim \frac{1}{(2\kappa+1)^{1/2}\left(1-\dfrac{z^2}{2\kappa+1}\right)^{1/4}}\begin{cases}\sin z(2\kappa+1)^{1/2}+O(\kappa^{-1/2}), \\ \qquad\qquad 0 \leqq |y| \leqq \epsilon, \\ \sin z(2\kappa+1)^{1/2}(1+O(\kappa^{-1/2})), \\ \qquad\qquad |y| > \epsilon,\end{cases}$$

in which

$$\arg(1 - z^2/(2\kappa+1)) < 0 \text{ when } y \geqq 0,$$

and ϵ is an arbitrary fixed positive quantity, sufficiently small.

THEOREM IV. *When the parameter κ is negative, the asymptotic forms of $w_j(z)$ inside the semi-circle $|z| = |2\kappa+1|^{1/2}$ contained in the half plane $\mathfrak{I}(z) \leqq 0$ are given by formulas* (27), *in which all restrictions on y are replaced by corresponding ones upon x.*

6. **The forms for bounded values of ξ.** The variable ξ is bounded in a region of $O(|\kappa|^{-2/3})$ about the origin of R_t, and here, as has been previously pointed out, the asymptotic forms of Theorems I to IV are not valid. For this range of values, the solutions $u_j(t)$ are described by special asymptotic formulas which are given in [L] as

$$u_j(t) = \Psi\xi^{1/3}J_{\mp 1/3}(\xi) + E(t,\kappa)/\kappa^{4/3}, \qquad j = 1, 2, \tag{28}$$

in which E is a bounded function, and J_α is the familiar symbol for the Bessel function of the first kind of order α. The function $\Psi = \Phi^{1/6}/\phi^{1/2}$, it may be noted, is analytic, single-valued, and non-vanishing in the neighborhood of $t=0$. The substitution of (28) into the identity (17) leads to asymptotic forms for $w_j(z)$ in which, by virtue of the fact that $t = O(|\kappa|^{-2/3})$ in the region under discussion, approximations for the functions ϕ, Φ and ξ readily obtainable from (13) may be used. The facts are embodied in the following theorem:

THEOREM V. *The asymptotic forms of $w_j(z)$ when $|\xi|$ is bounded are given by the formula*

$$(29)\ w_j(z) = \left(\frac{2\pi\Phi}{3\phi}\right)^{1/2} e^{\pi i/6}(2\kappa+1)^{1/2}\left\{C_{j1}^{(k)}J_{-1/3}(\xi) + C_{j2}^{(k)}J_{1/3}(\xi) + \frac{E(t,\kappa)}{\kappa}\right\}$$
$$(j = 1, 2;\ k = 2, 1),$$

in which $\arg\rho > 0$ *when* $k=2$, $\arg\rho \leqq 0$, $k=1$, *the approximations*

(30)
$$\left(\frac{\Phi}{\phi}\right)^{1/2} \sim \left(\frac{2}{3}\right)^{1/2}\left\{\frac{z}{(2\kappa+1)^{1/2}} - 1\right\}^{1/2},$$
$$\xi \sim \frac{i(2\kappa+1)2^{3/2}}{3}\left\{\frac{z}{(2\kappa+1)^{1/2}} - 1\right\}^{3/2}$$

may be used, and the coefficients $C_{jl}^{(k)}$ have the values shown in the table IV.

TABLE IV

k	$C_{11}^{(k)}$	$C_{12}^{(k)}$	$C_{21}^{(k)}$	$C_{22}^{(k)}$
2	$\left[\cos\frac{\pi}{2}(\kappa+\frac{2}{3})\right]$	$e^{2\pi i/3}\left[\cos\frac{\pi}{2}(\kappa-\frac{2}{3})\right]$	$\dfrac{\left[\sin\frac{\pi}{2}(\kappa+\frac{2}{3})\right]}{(2\kappa+1)^{1/2}}$	$\dfrac{e^{2\pi i/3}\left[\sin\frac{\pi}{2}(\kappa-\frac{2}{3})\right]}{(2\kappa+1)^{1/2}}$
1	$e^{-5\pi i/6}\left[\sin\frac{\pi}{2}(\kappa-\frac{2}{3})\right]$	$i\left[\sin\frac{\pi}{2}(\kappa+\frac{2}{3})\right]$	$\dfrac{e^{\pi i/6}\left[\cos\frac{\pi}{2}(\kappa-\frac{2}{3})\right]}{(2\kappa+1)^{1/2}}$	$\dfrac{-i\left[\cos\frac{\pi}{2}(\kappa+\frac{2}{3})\right]}{(2\kappa+1)^{1/2}}$

When ξ is small, formula (29) may be used to compute series which represent asymptotically the solutions w_i in the neighborhood of $z=(2\kappa+1)^{1/2}$. It is convenient here to introduce the variable η, defined by the relation

$$\eta = \frac{z^2 - 2\kappa - 1}{z^{2/3}}, \tag{31}$$

and since $z^2-2\kappa-1=o(|\kappa|^{1/3})$, when ξ is small, $\eta=o(1)$. Leading terms of the power series in η for the functions ϕ, ξ and ψ are readily computed, and are found to be the following:

$$\begin{aligned} \phi &= \frac{\eta^{1/2}}{z^{2/3}}\left(1+\frac{\eta}{2z^{4/3}}+\cdots\right), \\ \xi &= \frac{i}{3}\eta^{3/2}\left(1+\frac{\eta}{5z^{4/3}}+\cdots\right), \\ \frac{\Psi}{\rho^{1/3}} &= \frac{1}{e^{\pi i/6}3^{1/6}z^{2/3}}\left(1+\frac{17}{60z^{4/3}}\eta+\cdots\right). \end{aligned} \tag{32}$$

Since ξ is small, the quantities $\xi^{1/3}J_{\mp 1/3}(\xi)$ are essentially given by the leading terms of their power series expansions,

$$\begin{aligned} \xi^{1/3}J_{-1/3}(\xi) &= \frac{2^{1/3}}{\Gamma(\frac{2}{3})} - \frac{\xi^2}{2^{5/3}\Gamma(\frac{5}{3})} + \cdots, \\ \xi^{1/3}J_{1/3}(\xi) &= \frac{\xi^{2/3}}{2^{1/3}\Gamma(\frac{4}{3})} + \cdots, \end{aligned}$$

and these, with the use of (32) and the familiar formula

$$3^{1/2}\Gamma(\tfrac{1}{3})\Gamma(\tfrac{2}{3}) = 2\pi,$$

are expressible in terms of η as follows:

$$
\begin{aligned}
\xi^{1/3}J_{-1/3}(\xi) &= \frac{6^{1/3}3^{1/6}\Gamma(\frac{1}{3})}{2\pi}\left(1+\frac{\eta^3}{24}+\cdots\right),\\
\xi^{1/3}J_{1/3}(\xi) &= \frac{e^{\pi i/3}6^{2/3}3^{1/6}\Gamma(\frac{2}{3})}{4\pi}\eta+\cdots.
\end{aligned}
\tag{33}
$$

The quantities $(\Phi/\phi)^{1/2}J_{\mp 1/3}(\xi)$ found in (29) may also be written $\Psi(\xi/\rho)^{1/3}J_{\mp 1/3}(\xi)$; hence the leading terms of the expansions of these quantities in powers of η are readily determined from (32) and (33), and the asymptotic forms of $w_j(z)$ for small values of ξ are then at hand from the formula (29). The following theorem embodies the results:

THEOREM VI. *The asymptotic forms of $w_j(z)$ when $|\xi|$ is small are given by the series*

$$
\begin{aligned}
w_j(z) = \frac{(2\kappa+1)^{1/2}}{6^{1/6}\pi^{1/2}z^{2/3}}\Bigg\{ & D_{j1}^{(k)}\Gamma(\tfrac{1}{3}) + D_{j2}^{(k)}\frac{6^{1/3}\Gamma(\frac{2}{3})}{2}\eta \\
& + D_{j1}^{(k)}\frac{\Gamma(\frac{1}{3})}{24}\eta^3+\cdots+\frac{E(z,\kappa)}{\kappa}\Bigg\} \qquad (j=1,2;\ k=2,1),
\end{aligned}
\tag{34}
$$

in which $k=2$ when $-\pi/2<\arg(2\kappa+1)\leqq\pi/2$, $k=1$, when $-3\pi/2<\arg(2\kappa+1)\leqq-\pi/2$, η is defined by (31), *E is a bounded function, and the coefficients D are given in the table* V:

TABLE V

k	$D_{11}^{(k)}$	$D_{12}^{(k)}$	$D_{21}^{(k)}$	$D_{22}^{(k)}$
2	$\cos\frac{\pi}{2}(\kappa+\frac{2}{3})$	$-\cos\frac{\pi}{2}(\kappa-\frac{2}{3})$	$\dfrac{\sin\frac{\pi}{2}(\kappa+\frac{2}{3})}{(2\kappa+1)^{1/2}}$	$-\dfrac{\sin\frac{\pi}{2}(\kappa-\frac{2}{3})}{(2\kappa+1)^{1/2}}$
1	$e^{-5\pi i/6}\sin\frac{\pi}{2}(\kappa-\frac{2}{3})$	$i\sin\frac{\pi}{2}(\kappa+\frac{2}{3})$	$\dfrac{e^{\pi i/6}\cos\frac{\pi}{2}(\kappa-\frac{2}{3})}{(2\kappa+1)^{1/2}}$	$-\dfrac{i\cos\frac{\pi}{2}(\kappa+\frac{2}{3})}{(2\kappa+1)^{1/2}}$

7. **The asymptotic forms of the Hermite functions.** From the relation (4) between the equations (1) and (3), it is evident that the former equation has an even and an odd solution, and it is desirable to choose these in such a fashion that they reduce to the Hermite polynomials (2) when κ is a positive integer. This result is achieved if these solutions, here denoted as $U_{\kappa j}(z)$, be chosen so that

$$
U_{\kappa 1} = (-1)^{\kappa/2}\frac{\Gamma(\kappa+1)}{\Gamma\left(\frac{\kappa}{2}+1\right)}e^{z^2/2}w_1,\quad U_{\kappa 2} = 2(-1)^{(\kappa-1)/2}\frac{\Gamma(\kappa+1)}{\Gamma\left(\frac{\kappa+1}{2}\right)}e^{z^2/2}w_2,
\tag{35}
$$

the first of which reduces to the polynomials of even powers when κ is an even integer, the second to the polynomials of odd powers when κ is odd. The following theorem is then evident:

THEOREM VII. *The asymptotic forms of the solutions, $U_{\kappa j}(z)$, of the equation* (1), *which at $z=0$ have the values*

$$U_{\kappa 1}(0) = (-1)^{\kappa/2}\frac{\Gamma(\kappa+1)}{\Gamma\left(\frac{\kappa}{2}+1\right)}, \quad U'_{\kappa 1}(0) = 0,$$

$$U_{\kappa 2}(0) = 0, \quad U'_{\kappa 2}(0) = 2(-1)^{(\kappa-1)/2}\frac{\Gamma(\kappa+1)}{\Gamma\left(\frac{\kappa+1}{2}\right)},$$

are described by equations (35) *in terms of the asymptotic forms of $w_i(z)$, the latter, with their regions of validity, being given in Theorems* I *to* VI.

By restricting κ to positive integral values, the asymptotic forms of the Hermite polynomials (2) are obtained from (35). The simplification is obvious, except in the case of the coefficients of formula (21), $[\sin(\pi\kappa/2)]$, κ even, and $[\cos(\pi\kappa/2)]$, κ odd, the question being whether these reduce to quantities $[0]$, or vanish entirely. That the latter is the case becomes evident by noting, from (35), that $w_i \to 0$ as $x\to\infty$, but $e^{i\xi}\to 0$ and $e^{-i\xi}\to\infty$ as $x\to\infty$; hence, if the formula (21) for w_i is valid, the coefficients $[\sin(\pi\kappa/2)]$ and $[\cos(\pi\kappa/2)]$ must vanish when κ is respectively an even and an odd integer. With this fact, and the familiar asymptotic formula for the gamma function,

$$\Gamma(\kappa+1) \sim \kappa^{\kappa}e^{-\kappa}(2\pi\kappa)^{1/2}[1],$$

the following theorem is deduced from Theorem VII:

THEOREM VIII. *The asymptotic forms of the Hermite polynomials* (2) *are given by the formulas*

$$\text{(a)}\quad U_{\kappa}(z) \sim \frac{2^{\kappa/2}e^{(z^2+\kappa)/2}\kappa!}{\pi^{1/2}\kappa^{(\kappa+1)/2}\left(1-\dfrac{z^2}{2\kappa+1}\right)^{1/4}} \times\left\{\cos\left(\frac{\pi\kappa}{2}-z(2\kappa+1)^{1/2}\right)+O(\kappa^{-1/2})\right\},$$

$$\text{(b)}\quad U_{\kappa}(z) \sim \frac{2^{\kappa/2-1}e^{\kappa+z^2-z(z^2-2\kappa-1)^{1/2}/2}\kappa!}{\pi^{1/2}\kappa^{(\kappa+1)/2}\left(\dfrac{z^2}{2\kappa+1}-1\right)^{1/4}}$$

$$(36) \qquad \times\left\{\frac{z}{(2\kappa+1)^{1/2}}+\left(\frac{z^2}{2\kappa+1}-1\right)^{1/2}\right\}^{\kappa+1/2}[1],$$

$$\text{(c)} \quad U_\kappa(z) = \frac{e^{z^2/2}\kappa!}{(2\kappa+1)^{1/3}}\left(\frac{2e}{\kappa}\right)^{\kappa/2}\left(\frac{\Phi}{3\phi}\right)^{1/2}$$

$$\times\left\{e^{\pi i/6}J_{-1/3}(\xi)+e^{5\pi i/6}J_{1/3}(\xi)+O(\kappa^{-1})\right\},$$

$$\text{(d)} \quad U_\kappa(z) = \frac{e^{z^2/2}(2e)^{\kappa/2}\kappa!}{\pi^{1/2}\kappa^{\kappa/2}(2z)^{2/3}3^{1/6}}$$

$$\times\left\{\Gamma\left(\frac{1}{3}\right)-\frac{6^{1/3}\Gamma(\frac{2}{3})}{2}\eta+\frac{\Gamma(\frac{1}{3})}{24}\eta^3+\cdots+O(\kappa^{-1})\right\}$$

valid in the half plane $R(z) \geqq 0$ *as follows*:

(a) *inside the semi-circle* $|z| = (2\kappa+1)^{1/2}$;

(b) *in the regions* $\Xi^{(1)}$ *and* $\Xi^{(0)}$ *as shown in figure* 3(c);

(c) *in a region of dimension* $O(\kappa^{-1/6})$ *about the point* $z = (2\kappa+1)^{1/2}$;

(d) *for values of* z *in the neighborhood of the point* $z = (2\kappa+1)^{1/2}$.

The approximations (30) *may be used in* (c), *and the variable* η *of* (d) *is defined by* (31).

8. **Comparison with known formulas.** In the article to which reference has been made,* Watson determined the asymptotic forms of two independent solutions of the equation (3), $D_{-\kappa-1}(\pm iz)$, these formulas being

$$(37) \qquad D_{-\kappa-1}(\pm iz) = \frac{\pi e^{z^2/2}\left(\frac{\kappa}{e}\right)^{\kappa/2}}{\Gamma(\kappa+1)}e^{\mp iz(2\kappa)^{1/2}}\left\{e^{-z^2/2}+O(\kappa^{-1/2})\right\}$$

in which z is subject to the restriction,

$$|y| < \frac{2^{1/2}\kappa^\alpha}{6}, \quad 0 \leqq \alpha < \frac{1}{2},$$

and x is finite. The solutions are ordinarily represented by the power series

$$(38) \qquad D_{-\kappa-1}(\pm iz) = e^{z^2/2}\sum_{n=0}^{\infty}\frac{\Gamma\left(\frac{m+\kappa+1}{2}\right)}{m!\Gamma(\kappa+1)}(\mp 2iz)^m 2^{(\kappa-1)/2},$$

which are uniformly convergent in the finite part of the plane. Although Watson gave no bound for x in terms of κ, it is evident, because of the finiteness of x and the unboundedness of κ, that the values z may assume in (37) are included in the set of values z may take on in the formulas (27).

* Cf. Introduction, p. 340.

The solutions w_i are expressible by usual means in terms of $D_{-\kappa-1}(\pm iz)$, since the values of the latter and their derivatives at the origin are known from the series expansions (38). The computations show that

$$
\begin{aligned}
w_1(z) &= \frac{2^{-(\kappa+1)/2}\Gamma(\kappa+1)}{\Gamma\left(\dfrac{\kappa+1}{2}\right)}\{D_{-\kappa-1}(iz)+D_{-\kappa-1}(-iz)\},\\
w_2(z) &= \frac{i2^{-(\kappa+3)/2}\Gamma(\kappa+1)}{\Gamma\left(\dfrac{\kappa}{2}+1\right)}\{D_{-\kappa-1}(iz)-D_{-\kappa-1}(-iz)\},
\end{aligned}
\tag{39}
$$

and the substitution of the forms (37) into the equations (39), with the use of the asymptotic formula for the gamma function, leads to the forms

$$
w_1(z) \sim \begin{cases}\cos z(2\kappa)^{1/2} + O(\kappa^{-1/2}), & 0 \leqq |y| \leqq \epsilon,\\ \cos z(2\kappa)^{1/2}(1+O(\kappa^{-1/2})), & |y| > \epsilon,\end{cases}
$$

$$
w_2(z) \sim (2\kappa)^{-1/2}\begin{cases}\sin z(2\kappa)^{1/2} + O(\kappa^{-1/2}), & 0 \leqq |y| \leqq \epsilon,\\ \sin z(2\kappa)^{1/2}(1+O(\kappa^{-1/2})), & |y| > \epsilon,\end{cases}
$$

which are asymptotically equivalent to the formulas (27) when κ is large.

Adamoff's results for the Hermite polynomials on the finite part of the real axis are, of course, included in Watson's work and need not be considered separately. Plancherel and Rotach's results, however, take in the entire real axis and need separate consideration. Since they do not give a formula comparable to (36c), only the remaining formulas of (36) for $z=x$ need to be considered.

For the range of values $0 \leqq x < (2\kappa+2)^{1/2}$, they give the formula

$$
U_\kappa(x) \sim \frac{e^{(x^2+\kappa+1)/2}\kappa!2^{\kappa/2}}{\pi^{1/2}(\kappa+1)^{(\kappa+1)/2}\left(1-\dfrac{x^2}{2\kappa+2}\right)^{1/4}}\left\{\cos\left(\frac{\pi\kappa}{2}-x(2\kappa+2)^{1/2}\right)+O(\kappa^{-1/2})\right\},
$$

which is asymptotically equivalent to (36a) since

$$
(\kappa+1)^{(\kappa+1)/2} = \kappa^{(\kappa+1)/2}e^{1/2}[1].
$$

For $x > (2\kappa+2)^{1/2}$, the formula given by them may be reduced, by obvious simplifications, to the form

$$
\frac{2^{\kappa/2-1}\kappa!\exp\left(\dfrac{\kappa+x^2-x(x^2-2\kappa-2)^{1/2}}{2}\right)}{\pi^{1/2}\kappa^{(\kappa+1)/2}\left(\dfrac{x^2}{2\kappa+2}-1\right)^{1/4}}
$$

$$\times\left(\frac{x}{(2\kappa+2)^{1/2}}+\left\{\frac{x^2}{2\kappa+2}-1\right\}^{1/2}\right)^{\kappa+1/2}[1+O(\kappa^{-2/3})],$$

the ratio of which to (36b) is asymptotic to the quantity

$$\exp\left(-\frac{x}{2}\{(x^2-2\kappa-2)^{1/2}-(x^2-2\kappa-1)^{1/2}\}\right)\left(\frac{2\kappa+1}{2\kappa+2}\right)^{(2\kappa+1)/4};$$

the exponential in this expression is asymptotic to exp ($\frac{1}{4}$), and the second factor is asymptotic to exp ($-\frac{1}{4}$), so that the expression itself is asymptotic to unity. Hence the two forms under comparison are asymptotically equivalent.

In the neighborhood of the point $x=(2\kappa+2)^{1/2}$, Plancherel and Rotach determined the asymptotic form of the Hermite polynomials as a power series

$$(40)\qquad \frac{2^{\kappa-2/3}\kappa!e^{3x^2/4}}{\pi 3^{1/6}x^{\kappa+2/3}}\left\{\Gamma\left(\frac{1}{3}\right)+\frac{6^{1/3}\Gamma(\frac{2}{3})}{2^{4/3}}t-\frac{\Gamma(\frac{1}{3})}{6}t^3+\cdots+O(\kappa^{-1/3})\right\},$$

in which

$$t=\frac{2\kappa+2-x^2}{(2x)^{2/3}},$$

and, with a little manipulation, this series can be shown asymptotically equivalent to (37d). Now near $x=(2\kappa+1)^{1/2}$, $x-(2\kappa+1)^{1/2}=O(\kappa^{-1/6})$, so that

$$(2\kappa)^{1/2}=x\left(1+\frac{E_1}{\kappa^{2/3}}\right)$$

in which E_1 is a bounded function of κ. Hence

$$e^{\kappa/2}=\exp\left(\frac{x^2}{4}+E_1\kappa^{1/3}\right)[1+O(\kappa^{-1/3})],$$

$$(2\kappa)^{\kappa/2}=x^\kappa\left(1+\frac{E_1}{\kappa^{2/3}}\right)^\kappa;$$

but, since

$$\left(1+\frac{E}{n}\right)^n=e^E[1],$$

when n is large, it follows that

$$\left(1+\frac{E_1}{\kappa^{2/3}}\right)^\kappa=\left\{\left(1+\frac{E_1}{\kappa^{2/3}}\right)^{\kappa^{2/3}}\right\}^{\kappa^{1/3}}=\exp\{E_1\kappa^{1/3}(1+O(\kappa^{-2/3}))\},$$

or

$$\left(1 + \frac{E_1}{\kappa^{2/3}}\right)^\kappa = \exp\,(E_1\kappa^{1/3})[1 + O(\kappa^{-1/3})].$$

From these facts it is evident that

$$\frac{e^{\kappa/2}}{\kappa^{\kappa/2}} = \frac{e^{x^2/4}2^{\kappa/2}}{x^\kappa}[1 + O(\kappa^{-1/3})];$$

and since

$$\eta = -\,2^{2/3}t + O(\kappa^{-1/3}),$$

it is seen that the series (40) is equivalent asymptotically to the series (36d).

University of Wisconsin,
Madison, Wis.

CYCLOTOMY AND TRINOMIAL CONGRUENCES*

BY
L. E. DICKSON

1. Introduction. In the algebraic theory of cyclotomy we regard as known (or computed by tables of indices) one† or more of the functions $R(1, n)$, which Jacobi denoted by ψ_n. By rational operations and root extractions we obtain Jacobi's function F, then the periods, and finally the e^2 cyclotomic constants (k, h); see §6.

We here develop an arithmetical theory valid for every prime $p=ef+1$. The $R(1, n)$ are not computed by tables of indices, but are found in the simpler cases by representations of multiples of p by binary quadratic forms, and in general by a system of Diophantine equations (§§13–17). The cyclotomic constants (k, h) are found from linear equations, some of which are derived from linear relations between pairs of the functions $R(m, n)$. In an earlier memoir‡ we treated in full the cases $e=3, 4, 5, 6, 8, 10, 12$. Here we treat the cases in which e is a prime or a double of a prime. In particular, we find the number of solutions of $x^e+y^e\equiv\pm 1 \pmod p$.

2. Notations, results assumed. For proofs, see D.

Let g be a primitive root of a prime $p>2$. Let e be a divisor of $p-1=ef$. Let R be any root $\neq 1$ of $x^p=1$. The *periods* are

$$\eta_k = \sum_{t=0}^{f-1} R^{g^{et+k}} \qquad (k = 0, 1, \cdots, e-1). \tag{1}$$

Let (k, h) denote the number of sets of values of t and z, each chosen from $0, 1, \cdots, f-1$, for which

$$1 + g^{et+k} \equiv g^{ez+h} \pmod{p}. \tag{2}$$

$$\eta_m\eta_{m+k} = \sum_{h=0}^{e-1} (k, h)\eta_{m+h} + fn_k,\quad n_0 = 1\ (f \text{ even}),\quad n_{e/2} = 1\ (f \text{ odd}), \tag{3}$$

while the remaining n_k are zero.

$$\begin{gathered}(e-k, h-k) = (k, h),\\ (k, h) = (h, k)\ (f \text{ even}),\quad (k, h) = (h+\tfrac{1}{2}e, k+\tfrac{1}{2}e)\ (f \text{ odd});\end{gathered} \tag{4}$$

* Presented to the Society, December 28, 1934; received by the editors December 31, 1934.

† One function suffices when e is a prime <83, but not for $e=83$ or 107. See Kronecker, Journal für Mathematik, vol. 93 (1882), pp. 338–64; K. Schwering, ibid., vol. 102 (1888), p. 56, and Acta Mathematica, vol. 11 (1887–88), p. 119.

‡ American Journal of Mathematics, vol. 57 (1935), in press, cited as D.

$$\sum_{h=0}^{e-1}(k, h) = f - n_k \qquad (k = 0, 1, \cdots, e-1). \tag{5}$$

Let α be any $(p-1)$th root $\neq 1$ of unity. With Jacobi, write

$$F(\alpha) = \sum_{k=0}^{p-2} \alpha^k R^{g^k}. \tag{6}$$

In particular, let $\alpha = \beta^m$, where β is a primitive eth root of unity. Then

$$F(\beta^m) = \sum_{j=0}^{e-1} \beta^{mj}\eta_j, \qquad F(\beta^n)F(\beta^{-n}) = (-1)^{nf}p, \tag{7}$$

if n is not divisible by e. Since $R^{p-1} + \cdots + R + 1 = 0$ is irreducible* in the field of the rational functions with integral coefficients of a primitive kth root of unity when k is not divisible by p, it follows that $F(\alpha) \neq 0$.

If no one of m, n, $m+n$ is divisible by e,

$$R(m, n, \beta) \equiv \frac{F(\beta^m)F(\beta^n)}{F(\beta^{m+n})} = \sum_{k=0}^{e-1} \beta^{nk} \sum_{h=0}^{e-1} \beta^{-(m+n)h}(k, h), \tag{8}$$

$$R(n, m) = R(m, n) = (-1)^{nf}R(-m-n, n), \tag{9}$$

$$R(m, n)R(-m, -n) = p. \tag{10}$$

$$R(m, n, \beta^j) = R(jm, jn, \beta). \tag{11}$$

$$R(2r, 2s, \beta)_e = R(r, s, \beta^2)_E, \text{ if } e = 2E; \tag{12}$$

$$(k, h)_E = (k, h) + (k + E, h) + (k, E + h) + (E - k, h - k). \tag{13}$$

$$F(-1)F(\alpha^2) = \alpha^{2m}F(\alpha)F(-\alpha), \; g^m \equiv 2 \pmod{p}. \tag{14}$$

When g is replaced by a new primitive root g^r of p,

$$R(m, n) \text{ becomes } R(mr', nr'), \; r'r \equiv 1 \pmod{e}. \tag{15}$$

3. **Relations between the coefficients of $R(1, n)$.** We employ (9) for $m=1$ and then (8) for $m = -1-n$. The product of the two powers of β now has the exponent $nk+h$. Eliminate h by use of $nk+h \equiv i \pmod{e}$. We get

$$(-1)^{nf}R(1, n) = \sum_{i=0}^{e-1} B(i, n)\beta^i, \qquad B(i, n) = \sum_{k=0}^{e-1}(k, i - nk). \tag{16}$$

When n and k are fixed, $i - nk$ ranges with i over a complete set of residues modulo e. Hence by (5),

$$\sum_{k,i}(k, i - nk) = \sum_{k,h}(k, h) = \sum_k (f - n_k) = ef - 1,$$

* Kronecker, Journal de Mathématiques, vol. 19 (1854), pp. 177–92.

since a single n_k is 1 and the others are 0. Hence

$$\sum_{i=0}^{e-1} B(i, n) = p - 2. \tag{17}$$

THEOREM* 1. *If e is prime*† *to* 6,

$$\sum_{i=0}^{e-1} iB(i, n) \equiv 0, \qquad \sum_{i=0}^{e-1} i^2 B(i, n) \equiv 0 \qquad (\mathrm{mod}\ e). \tag{18}$$

Since $p = ef+1 > 2$, f is even. Write $B(i)$ for $B(i, n)$. By (4),

$$(0, r), \quad (r, 0), \quad (-r, -r), \quad r \not\equiv 0 \ (\mathrm{mod}\ e), \tag{19}$$

are equal. They are respectively terms of $B(r)$ with $k=0$, $B(n, r)$ with $k=r$, $B(-r, -nr)$ with $k=-r$. Since the sum of the three arguments of B is zero, the sum of the corresponding terms of (18_1) is zero.

Next let $r \not\equiv 0$, $s \not\equiv 0$, $r \not\equiv s$ (mod e). Then each of

$$(r, s) = (-r, s-r) = (-s, r-s) = (s, r) = (s-r, -r) = (r-s, -s) \tag{20}$$

has incongruent arguments each $\not\equiv 0$ (mod e); while no two have congruent first arguments and congruent second arguments. Hence these (r, s) coincide in sets of six. They are terms of

$$B(s+nr), \quad B(s-r-nr), \quad B(r-s-ns), \quad B(r+ns),$$

$$B(-r+ns-nr), \quad B(-s+nr-ns),$$

the sum of whose arguments is zero. This proves (18_1).

In (18_2), the sum of the coefficients of the three numbers (19) and the sum of the coefficients of the six numbers (20) are respectively

$$(1+n+n^2)r^2, \quad 2(1+n+n^2)(r^2 - rs + s^2).$$

Multiply (5) by k^2 and sum for k. Thus

$$\sum_{k,h=0}^{e-1} k^2(k, h) = fN, \qquad N = \sum_{k=1}^{e-1} k^2 = \tfrac{1}{6}e(e-1)(2e-1).$$

Since $e \equiv \pm 1$ (mod 6), N is a multiple of e. In the double sum the sum of the coefficients of the three numbers (19) or six numbers (20) is respectively

$$0 + r^2 + r^2, \quad r^2 + r^2 + s^2 + s^2 + 2(s-r)^2 = 4(r^2 - rs + s^2).$$

---

* The case in which $n=1$ and e is a prime >3 is due to K. Schwering, Journal für Mathematik, vol. 93 (1882), pp. 334–37. His formulas have $-2rs$ in error for $-rs$. By (17) and Theorem 1, $1+R(1, n)$ is divisible by $(1-\beta)^3$. Our new application of Theorem 1 is given in §13.

† When e is composite (16) is not to be reduced to degree $<e-1$ by the equation satisfied by β.

THEOREM 2. *If e is prime to 2, 3 and 5,*

$$\sum_{i=0}^{e-1} i^4 B(i, n) \equiv 0 \qquad (\text{mod } e).$$

The sum of the coefficients of the three numbers (19) and the sum of the coefficients of the six numbers (20) are respectively

$$2r^4J, \qquad 2\{r^4 + s^4 + (r-s)^4\}J, \qquad J = 1 + 2n + 3n^2 + 2n^3 + n^4.$$

But $2r^4$ and $2\{r^4+s^4+(r-s)^4\}$ are evidently the corresponding sums in

$$\sum_{k,h=0}^{e-1} k^4(k, h) = fM, \qquad M = \sum_{k=1}^{e-1} k^4 = (e-1)e(2e-1)(3e^2-3e-1)/30.$$

The final factor is a multiple of 5 if $e \equiv -1$ or $2 (\text{mod } 5)$. Also, $2e-1 \equiv 0$ if $e \equiv -2$ (mod 5). Hence M is a multiple of e if e is prime to 5 and 6.

Since $1+\beta+\cdots+\beta^{e-1}=0$, we may eliminate the constant term from (16) and obtain (for e prime to 6)

$$(21) \qquad R(1, n, \beta) = \sum_{i=1}^{e-1} a_i\beta^i, \qquad a_i = B(i, n) - B(0, n),$$

$$(22) \qquad \sum_{i=1}^{e-1} a_i = p - 2 - e\sum_{k=0}^{e-1} (k, -nk) \equiv -1 \qquad (\text{mod } e),$$

$$(23) \qquad \sum_{i=1}^{e-1} ia_i \equiv 0, \qquad \sum_{i=1}^{e-1} i^2a_i \equiv 0, \qquad \sum_{i=1}^{e-1} i^4a_i \equiv 0 \qquad (\text{mod } e),$$

where e is prime to 30 for the third. Every linear homogeneous function of the a_i which is a multiple of e for all consistent values of the (k, h) is a linear combination of the three in (23), at least when e is a prime.

Write $e=2E+1$. In the terms $i=E+1, \cdots, 2E$ write $I=e-i$. Then the last two of (23) give

$$(24) \qquad \sum_{i=1}^{E} i^2b_i \equiv 0, \qquad \sum_{i=1}^{E} i^4b_i \equiv 0 \ (\text{mod } e), \qquad b_i = a_i + a_{e-i},$$

when e is prime to 3 or 15, respectively.

4. **To express p as a sum of multiples of squares.** By (10), (11), and (21),

$$p = R(1, n, \beta)R(1, n, \beta^{-1}) = \sum a_i^2 + \sum_{i=1}^{e-2} (\beta^i + \beta^{-i}) \sum_{j=1}^{e-1-i} a_ja_{j+i}.$$

In the terms with $i \geqq E+1$, write $i=e-I$. Hence if $e=2E+1$,

$$(25) \qquad p - \sum_{i=1}^{e-1} a_i^2 = \sum_{i=1}^{E} (\beta^i + \beta^{-i})C_i, \qquad C_i = \sum_{j=1}^{e-1-i} a_ja_{j+i} + \sum_{j=1}^{i-1} a_ja_{j+e-i},$$

where the final sum is absent if $i=1$.

Let e be an odd prime. Then $\beta^{e-1}+\cdots+\beta+1$ is irreducible in the field of rational numbers. Thus $C_1, \cdots, C_E$ are equal, and

$$p = \sum_{i=1}^{e-1} a_i^2 - C/E, \qquad C = \sum_{i=1}^{E} C_i = \sum a_1 a_2, \tag{26}$$

where the final sum is a symmetric function. Next

$$(e-1)^2\left\{\sum_{i=1}^{e-1} a_i^2 - E^{-1}\sum a_1 a_2\right\} - \left(\sum_{i=1}^{e-1} a_i\right)^2 = eM, \tag{27}$$

$$M = (e-2)\sum_{i=1}^{e-1} a_i^2 - 2\sum a_1 a_2, \qquad D = \sum_{i=1}^{E} (a_i - a_{e-i})^2, \tag{28}$$

$$M - ED = \Delta = (E-1)\sum_{i=1}^{E} b_i^2 - 2\sum b_1 b_2, \tag{29}$$

where the final sum is symmetric in $b_1, \cdots, b_E$ defined in (24). The proof of (29) follows from

$$2\sum b_1 b_2 = \sum b_i b_j = \sum (a_i a_j + a_{e-i} a_{e-j} + 2a_i a_{e-j}), \quad i \neq j.$$

THEOREM 3. $(e-1)^2 p = (\sum a_i)^2 + e(ED+\Delta)$.

Case $e=5$. Then $\Delta = (b_1-b_2)^2$. By (24), $b_1 - b_2 = 5W$, where W is an integer. Hence

$$16p = (\sum a_i)^2 + 125W^2 + 10(a_1-a_4)^2 + 10(a_2-a_3)^2. \tag{30}$$

Denote (29) by Δ_E and consider the like function Δ_j of $b_1, \cdots, b_j$. Define

$$L_{j-1} = (j-1)b_j - \sum_{i=1}^{j-1} b_i. \tag{31}$$

The recursion formula

$$(j-1)\Delta_j = j\Delta_{j-1} + L_{j-1}^2, \quad \Delta_1 = 0, \tag{32}$$

yields Δ_j as a linear combination of $L_{j-1}^2, \cdots, L_1^2$:

$$\Delta_2 = L_1^2, \quad 2\Delta_3 = L_2^2 + 3L_1^2, \quad 3\Delta_4 = L_3^2 + 2L_2^2 + 6L_1^2,$$
$$12\Delta_5 = 3L_4^2 + 5L_3^2 + 10L_2^2 + 30L_1^2, \quad 10\Delta_6 = 2L_5^2 + 12\Delta_5.$$

Case $e=7$. Then (24) give $b_1 \equiv b_2 \equiv b_3 \pmod 7$. Thus

$$\begin{gathered} L_1 = b_2 - b_1 = 7v, \quad L_2 = 2b_3 - b_1 - b_2 = 7W, \\ 72p = 2(\sum a_i)^2 + 42D + 7^3(W^2+3v^2), \\ D = (a_1-a_6)^2 + (a_2-a_5)^2 + (a_3-a_4)^2. \end{gathered} \tag{33}$$

Case $e=11$. Then

$$(34)\quad 1200p = 12(\sum a_i)^2 + 660D + 11(3L_4^2 + 5L_3^2 + 10L_2^2 + 30L_1^2).$$

By (24),

$$(35)\quad b_4 \equiv -2b_1 + 4b_2 - b_3, \qquad b_5 \equiv 3b_1 + 3b_2 - 5b_3 \pmod{11},$$

$$(36)\quad L_3 + 2L_2 + 2L_1 \equiv 0,\ L_4 - L_2 + 3L_1 \equiv 0 \pmod{11}.$$

But it is not possible to segregate from Δ_5 a square divisible by 11, unlike the cases $e=5$, $e=7$.

5. **Sets of conjugate** $R(1, j)$. If j is prime to e and if β is replaced by β^j, the coefficients a_i in (21) are permuted cyclically, whence $\sum a_i$ is unaltered. For example, if $e=5$ and $j=3$, the substitution is $(a_1\ a_2\ a_4\ a_3)$. By (11), $R(m, n)$ becomes $R(jm, jn)$, which will be called a *conjugate* to $R(m, n)$. Hence each $R(m, n)$ is conjugate to some $R(1, -)$.

Let $e=2E+1$ be a prime >3. By (9), $R(1, 1)$ equals $R(1, e-2)$, which is conjugate to $R(E, 1)=R(1, E)$ since $E(e-2)\equiv 1 \pmod{e}$.

Consider a new $R(1, n)$, where $2\leqq n\leqq e-3$, $n\neq E$. It equals $R(1, e-1-n)$, which is conjugate to $R(m, 1)=R(1, e-1-m)$ if $m(1+n)\equiv -1 \pmod{e}$. Again, $R(1, n)$ equals $R(n, e-1-n)$, which is conjugate to $R(1, e-1-t)$ $=R(1, t)$ if $tn\equiv 1 \pmod{e}$. We now have six $R(1, -)$ whose second arguments are congruent modulo e to

$$n,\ 1/n,\ -1-n,\ -1-1/n,\ -1/(n+1),\ -n/(n+1).$$

These form a group (of cross ratios). Hence they are distinct unless

$$(37)\quad n^2+n+1\equiv 0,\quad (2n+1)^2\equiv -3 \pmod{e},\qquad e\equiv 1 \pmod{3}.$$

THEOREM 4. *If e is a prime >3, $R(1, 1), \cdots, R(1, e-2)$ fall into complete sets of conjugates as follows: when $e=6r+5$, one set of three, and r sets of six; when $e=6r+1$, one set of three, $r-1$ sets of six, and the set*

$$(38)\quad R(1, n),\ R(1, e-1-n),\ n^2+n+1\equiv 0 \pmod{e}.$$

6. **Determination of the e^2 cyclotomic constants (k, h) when e is a prime** >3. For $n\leqq e-2$, let $R(1, n)$ be unaltered when β is replaced by β^j, $1<j<e$. By (11), $R(1, n)=R(j, jn)$. By (9), one of j, jn must be congruent modulo e to 1 or n. If $jn\equiv 1$, either $j=n$, or $-1-j\equiv n$ and $n^2+n+1\equiv 0$. Next, let $j=n$. If $n^2\equiv 1$, then $n\equiv 1$, $j\equiv 1$, contrary to hypothesis. Hence $-n-n^2\equiv 1$ $(\text{mod } e)$. Hence $R(1, n)$ is unaltered if, and only if, $j=n$, $n^2+n+1\equiv 0$ (mod e). Then $e\equiv 1 \pmod{6}$ by (37). Since $n^3\equiv 1 \pmod{e}$, the substitution which replaces β by β^n is of period 3. Hence for (38) there are only $\frac{1}{3}(e-1)$ distinct a_i in R, while any R except those two has $e-1$ distinct a_i.

The linear relations (5) with $k \geqq \frac{1}{2}(e+1)$ reduce by (4) to those with $k \leqq \frac{1}{2}(e-1)$, whence there are only $\frac{1}{2}(e+1)$ independent relations. Then Theorem 4 shows that the sum of all the a's in the $R(1, n)$, one from each set of conjugates, increased by the preceding $\frac{1}{2}(e+1)$, is $\frac{1}{6}(e+1)(e+2)$ whether $e \equiv 5$ or 1 (mod 6). But this is the number of reduced (i, j) which remain after deleting duplicates by (4). In proof, we retain $(0, 0)$ and $(0, r)$ for $r=1, \cdots, e-1$. Each of the latter is one of a set of three equal (i, j) by (19). The remaining $e^2-3(e-1)-1$ fall into sets of six by (20). Hence the number of reduced (i, j) is

$$1 + e - 1 + \tfrac{1}{6}(e^2 - 3e + 2) = \tfrac{1}{6}(e + 1)(e + 2).$$

These $\frac{1}{6}(e+1)(e+2)$ linear equations in the same number of unknowns are linearly independent and hence determine the unknowns uniquely, and therefore determine all the cyclotomic constants. First, we note that (8) and (7_2) imply

$$[F(\beta)]^e = pR(1, 1)R(1, 2) \cdots R(1, e - 2). \tag{39}$$

Hence the $R(1, n)$ and their conjugates determine the $F(\beta^m)$, $m=0, \cdots, e-1$. By (7_1), the latter determine the periods η_j. Then (3) determine $(k, 0)$, $(k, 1), \cdots, (k, e-1)$ for $k=0, \cdots, e-1$, since the determinant of their coefficients is the cyclic determinant having $\eta_0, \cdots, \eta_{e-1}$ in the first row and hence is the product of the linear functions (7_1) whose values are the $F(\beta^m)$. But the latter were proved in §2 to be not zero.

THEOREM 5. *Let e be a prime >3. The e^2 numbers (i, j) reduce to $M=\frac{1}{6}(e+1)\cdot(e+2)$ after deleting duplicates by* (4). *These M reduced (i, j) are uniquely determined by M linear equations composed of* (5) *for $k=0, \cdots, \frac{1}{2}(e-1)$, and $M-\frac{1}{2}(e+1)$ equations expressing the a_{in} in terms of the (k, h), where there are $\frac{1}{3}(e-1)$ distinct a_{in} in case* (38), *and $e-1$ a_{in} for the further $R(1, n)$ in Theorem* 4, *one from each set of conjugates.*

For example, if $e=5$, $M=7$, $M-\frac{1}{2}(e+1)=4$, and we employ only the four a_{i1}.

We shall give details only for (0, 0). By D, §10, either of

$$x^e + y^e \equiv \pm 1 \qquad (\mathrm{mod}\ p = ef + 1),\ f \text{ even}, \tag{40}$$

has exactly $2e+e^2$ (0, 0) solutions.

7. **To find (0, 0).** Employ a second subscript n in (22). Then

$$\sum_{n=1}^{e-2}\sum_{i=1}^{e-1} a_{in} = (e - 2)(p - 2) - e(e - 2)(0, 0) - eS, \quad S = \sum_{k=1}^{e-1}\sum_{n=1}^{e-2} (k, - nk).$$

Let e be an odd prime. Each k is not divisible by e. For a fixed k, and $n=1, \cdots, e-2$, the residues modulo e of $-nk$ are $1, \cdots, e-1$, except k. Hence

$$S = \sum_{k=1}^{e-1} \left\{ \sum_{h=1}^{e-1} (k, h) - (k, k) \right\}.$$

We may extend the summation from $h=0$ by subtracting the new term $(k, 0)$. But $(k, k)=(-k, 0)$ and $-k$ ranges with k over $1, \cdots, e-1$ modulo e. Thus

$$S = \sum_{k=1}^{e-1} \sum_{h=0}^{e-1} (k, h) - 2\sum_{h=1}^{e-1} (0, h) = (e-1)f - 2\{f - 1 - (0, 0)\}$$

by (5). From eS eliminate $ef=p-1$. Hence

$$\sum_{n=1}^{e-2} \sum_{i=1}^{e-1} a_{in} = p - 3e + 1 - e^2(0, 0). \tag{41}$$

This determines $(0, 0)$. The left member may be simplified by Theorem 4. If $e=5$ we obtain, in accord with (66) of D,

$$25(0, 0) = p - 14 - 3\sum_{i=1}^{4} a_{i1}; \tag{42}$$

$$3\sum_{1}^{6} a_{i1} + 2\sum_{1}^{6} a_{i2} = p - 20 - 49(0, 0) \text{ if } e = 7. \tag{43}$$

In Theorem 3, $D \geqq 0$, $\Delta \geqq 0$, and p is not a square. Hence

$$(e-1)p^{1/2} > \sum_{i=1}^{e-1} a_{in}. \tag{44}$$

This with (41) gives*

$$e^2(0, 0) > p - 3e + 1 - (e-1)(e-2)p^{1/2}. \tag{45}$$

Cyclotomy for $e=2E$, E an odd prime, §§8–12

8. **Sets of conjugate** $R(i, j)$. If m and n are both even,

$$R(2m, 2n) = R(e - 2m - 2n, 2n),$$

and the latter first argument is double an odd integer. Hence we may assume that m is odd and prime to E. Thus $mx \equiv 1 \pmod{E}$ determines x, and $R(2m, 2n)$ is conjugate to $R(2, 2nx)$. The latter is found by (12). The dis-

* The writer gave a different proof in Journal für Mathematik, vol. 135 (1909), pp. 181–88. At bottom of p. 188, insert $p \neq 41$.

tribution of the $R(2, \text{even})$ into conjugate sets is therefore obtained from Theorem 4 applied to E in place of e, with subsequent multiplication of arguments by 2.

Consider therefore $R(m, n)$, where m is odd. If $m\neq E$, $mx\equiv 1 \pmod e$ is solvable for x. Next, $R(E, k)$ is excluded if k is a multiple of E. If k is even, pass to the equal $R(-E-k, k)$, whose first argument is odd and prime to e. Hence any R is conjugate to an $R(1, -)$. We extend the definition of conjugate so that $(-1)^f R$ is conjugate to R.

THEOREM 6. *If $\frac{1}{2}e=E$ is an odd prime, $R(1, 1), \cdots, R(1, e-2)$ fall into $\frac{1}{2}(E+1)$ sets of conjugates as follows:*

$$R(1, 1) = (-1)^f R(1, e-2);$$
$$R(1, E) = (-1)^f R(1, E-1);$$
$$R(1, j) = (-1)^f R(1, e-1-j), \quad R(1, x) = (-1)^f R(1, e-1-x),$$
$$1 < j < E, \quad j \text{ odd}, \quad xj \equiv 1 \pmod e.$$

9. **To find (0, 0) and (0, E).** We have

$$\beta^E = -1, \quad \beta^{E-1} - \cdots - \beta + 1 = 0,$$

$$R(1, n) = \sum_{i=0}^{E-1} r_i\beta^i = \sum_{i=1}^{E-1} R_i\beta^i, \quad R_i = r_i - (-1)^i r_0,$$

$$\sum_{i=1}^{E-1} (-1)^i R_i = \sigma - Er_0, \quad \sigma = \sum_{i=0}^{E-1} (-1)^i r_i = R(1, n) \text{ for } \beta = -1,$$

$$R(1, n) = \sum_{k=0}^{e-1} \beta^{nk} Y(k, \beta), \quad Y(k, \beta) = \sum_{h=0}^{e-1} \beta^{-(1+n)k}(k, h),$$

by (8). Henceforth, let n be odd.* Then by (5)

$$Y(k, -1) = \sum_{h=0}^{e-1} (k, h) = f - n_k, \quad \sigma = \sum_{k=0}^{e-1} (-1)^k (f - n_k) = -\sum_{k=0}^{e-1} (-1)^k n_k,$$

whence $\sigma = -(-1)^f$. We get r_0 from (16) by using the terms with $i=0$ and $i=E$.

THEOREM 7. *If $\frac{1}{2}e=E$ and n are odd,*

$$R(1, n) = \sum_{i=1}^{E-1} R_i\beta^i, \quad \rho(1, n) \equiv \sum_{i=1}^{E-1} (-1)^i R_i = -(-1)^f\{1 + Er(1, n)\},$$

$$r(1, n) = \sum_{k=0}^{e-1} \{(k, -nk) - (k, E - nk)\}.$$

* If n were even, pass to $R(1, -1-n)$ by (9).

If $\frac{1}{2}e=E$ is an odd prime, we find that

$$\rho(1,1)+2\sum\rho(1,n)=-(-1)^f\{E+E^2(0,0)-E^2(0,E)\}, \tag{46}$$

where $\rho(1, 1)$ and the $\rho(1, n)$ together correspond to the $R(1, k)$ with odd k's such that a single one is chosen from each set of conjugates in Theorem 6. For example, the $k>1$ are

3 if $e=6$; 5, 7 if $e=10$; 3, 7, 9 if $e=14$; 3, 5, 7, 11, 13 if $e=22$.

Special cases of (13) are

$$(0,0)_E=(0,0)+3(0,E),\ f \text{ even};\quad (0,0)_E=3(0,0)+(0,E),\ f \text{ odd}. \tag{47}$$

By this and (46) we get (0, 0) and (0, E) in terms of the $\rho(1, k)$. To find the latter we shall next evaluate the $R(1, k)$.

10. **Determination of the** $R(1, k)$. By (14) for $\alpha=\beta^j$,

$$\begin{aligned} F(\beta^E)F(\beta^{2j})&=\beta^{2mj}F(\beta^j)F(\beta^{j+E}),\\ R(E,2j)&=\beta^{2mj}R(j,j+E). \end{aligned} \tag{48}$$

Multiply the former by $F(\beta^j)/(F(\beta^{j+E})F(\beta^{2j}))$. Thus

$$R(j,E)=\beta^{2mj}R(j,j),\quad R(1,E)=\beta^{2m}R(1,1). \tag{49}$$

By (48) for $j=\frac{1}{2}(E-1)$ and (9),

$$R(1,E-1)=(-1)^{f(E-1)/2}\beta^{m(E-1)}R(1,\tfrac{1}{2}(E-1)). \tag{50}$$

By (49_2) and (9),

$$R(1,E-1)=(-1)^f\beta^{2m}R(1,1),\quad F(\beta^2)F(\beta^{E-1})=(-1)^f\beta^{2m}F(\beta)F(\beta^E), \tag{51}$$

$$R(2,E-1)=(-1)^f\beta^{2m}R(1,E)=(-1)^f\beta^{4m}R(1,1), \tag{52}$$

by (49). By (50) and (51_1),

$$R(1,\tfrac{1}{2}(E-1))=(-1)^{f(E+1)/2}\beta^{m(3-E)}R(1,1). \tag{53}$$

By (12), $R(2, 2)=R(1, 1)_E$ is known. Replacing β by $\beta^{(E-1)/2}$, we get $R(E-1, E-1)=R(2, E-1)$. Then by (52), (53), (51), (49_2), we get $R(1, 1)$, $R(1, \frac{1}{2}(E-1))$, $R(1, E-1)$, $R(1, E)$.

If $e=6$, we have the desired $R(1, 1)$, $R(1, 3)$. If $e=10$, we have the desired $R(1, 1)$, $R(1, 5)$, $R(1, 7)$. By (8),

$$R(m,t)R(n,m+t)=R(m,n)R(m+n,t). \tag{54}$$

For $E>3$, take $m=n=1$, $t=(E-3)/2$. Thus

$$R(1,\tfrac{1}{2}(E-3))R(1,\tfrac{1}{2}(E-1))=R(1,1)R(2,\tfrac{1}{2}(E-3)). \tag{55}$$

By (53) this gives

$$R(1, 2) = \beta^{4m}R(2, 2) \text{ if } e = 14;\; R(1, 4) = \beta^{8m}R(2, 4) \text{ if } e = 22. \tag{56}$$

For $e=14$, replace β by β^9 in $R(1, 11)=(-1)^f R(1, 2)$; we get $R(1, 9)$. Hence we have the desired $R(1, 1)$, $R(1, 3)$, $R(1, 7)$, $R(1, 9)$.

For $e=22$, replace β by β^{13} in $R(1, 17)=(-1)^f R(1, 4)$; we get $R(1, 13)$. We have also $R(1, 1)$, $R(1, 5)$, $R(1, 11)$. We desire also $R(1, 3)$ and $R(1, 7)$. By (54) for $m=1$, $t=2$, $n=5$,

$$R(1, 2)R(5, 3) = R(1, 5)R(6, 2),\quad R(3, 5) = R(1, 5, \beta^5), \tag{57}$$

which give $R(1, 2)$. Replacing β by β^7 in $R(1, 19)=(-1)^f R(1, 2)$, we get $R(1, 7)$. By (54) for $m=n=1$, $t=2$,

$$R(1, 2)R(1, 3) = R(1, 1)R(2, 2), \tag{58}$$

which gives $R(1, 3)$. But $R(1, 7)$ and $R(1, 3)$ were not found linearly. Neither is a product of a unit by any R not conjugate to itself.

11. **Theory*** for $e=14$. We have

$$\beta^7 = -1,\quad \beta^6 - \beta^5 + \beta^4 - \beta^3 + \beta^2 - \beta + 1 = 0. \tag{59}$$

Thus $B=\beta^2$ is a primitive seventh root of unity. We may regard as known (§14) the a_i in

$$R(1, 1)_{E=7} = a_1B + \cdots + a_6B^6. \tag{60}$$

Replacing B by B^3 (and reducing by $B^7=1$), we get $R(3, 3)$ for $E=7$. We now write β^2 for B, reduce by (59_1), and by (12) get

$$R(4, 4) = -a_2\beta + a_4\beta^2 - a_6\beta^3 + a_1\beta^4 - a_3\beta^5 + a_5\beta^6,$$
$$R(2, 6) = R(6, 6) = -a_6\beta + a_5\beta^2 - a_4\beta^3 + a_3\beta^4 - a_2\beta^5 + a_1\beta^6.$$

By (52), (53), (49_1) and the remark below (56),

$$\begin{aligned} R(1, 1) &= (-1)^f\beta^{-4m}R(2, 6),\quad R(1, 3) = (-1)^f\beta^{-8m}R(2, 6),\\ R(1, 7) &= (-1)^f\beta^{-2m}R(2, 6),\quad R(1, 9) = (-1)^f\beta^{8m}R(4, 4). \end{aligned} \tag{61}$$

By (46)

$$\begin{aligned} &\rho(1, 1) + 2\rho(1, 3) + 2\rho(1, 7) + 2\rho(1, 9)\\ &\quad = -(-1)^f\{7 + 49(0, 0) - 49(0, 7)\}. \end{aligned} \tag{62}$$

By Theorem 7, $\rho(1, n)$ is derived at once from $R(1, n)$.

I. Let $m\equiv 0 \pmod 7$, i.e., 2 is a residue of a seventh power modulo p. For $p<1000$ this case occurs only when $p=631$, 673, 953. Then

$$\rho(1, j) = (-1)^f\sum_1^6 a_i \text{ if } j = 1, 3, 7 \text{ or } 9.$$

* Details for $e=6$ and $e=10$ were given in D.

Cancelling $7(-1)^f$ from (62), we get

$$- 1 - 7(0, 0) + 7(0, 7) = \sum a_i. \tag{63}$$

This with (47) yields at once (0, 0) and (0, 7).

II. $m \equiv 4 \pmod 7$; true for $p<1000$ only when $p=29, 43, 127, 421, 701, 967$. We get

$$\rho(1, j) = (-1)^f\{\sum a_i - 7a_k\},$$

$$k = 5 \text{ if } j = 1,\ k = 3 \text{ if } j = 3,\ k = 6 \text{ if } j = 7 \text{ or } 9;$$

$$7(0, 0) - 7(0, 7) = - 1 - a_1 - a_2 + a_3 - a_4 + 3a_6. \tag{64}$$

III. $m \equiv 2 \pmod 7$. Thus $p=71, 281, 449, 547, 659, 743, 911$, etc. Now $k=6$ if $j=1$, $k=5$ if $j=3$, $k=3$ if $j=7$ or 9,

$$7(0, 0) - 7(0, 7) = - 1 - a_1 - a_2 + 3a_3 - a_4 + a_5. \tag{65}$$

This and the further cases are similar to II. For, if we use a new primitive root G, where $g \equiv G^t \pmod p$, m is replaced by $M \equiv mt \pmod{p-1}$. We can choose t so that $M \equiv 0$ or $4 \pmod 7$.

12. **Determination of all** (k, h). Let $e=2E$, E a prime >3. There remain $m=(e+1)(e+2)/6$ reduced (i, j) after deleting duplicates by (4). There are $M=(E+1)(E+2)/6$ reduced $(k, h)_E$, which we regard as known by the theory for E. To these correspond M linear equations (13); the first terms (k, h) of their second members may be eliminated from the $E+1$ independent relations (5) for $k=0, 1, \cdots, E$. We must evidently get the $\frac{1}{2}(E+1)$ independent relations (5) with $0 \leqq k \leqq \frac{1}{2}(E-1)$ involving the $(k, h)_E$. We discard these relations between knowns, and retain only the $\frac{1}{2}(E+1)$ new relations.

By Theorem 6 the $R(1, n)$ form $\frac{1}{2}(E+1)$ sets of conjugates. Each $R(1, n)$ is a linear combination of $\beta, \cdots, \beta^{E-1}$. Retaining only one R from each set, we have $\frac{1}{2}(E+1)(E-1)$ coefficients. We now have

$$M + \tfrac{1}{2}(E + 1) + \tfrac{1}{2}(E + 1)(E - 1) = m$$

linear equations for the m reduced (i, j). These equations are independent by the discussion following (39).

Diophantine equations in the a_j, §§13–17

13. **General theory.** Let e be an odd prime. With the abbreviation (25) for the quadratic functions C_i of the a_j, we have the system of $E=\frac{1}{2}(e-1)$ quadratic Diophantine equations

$$C_1 = C_2 = \cdots = C_E, \qquad p = \sum_{j=1}^{e-1} a_j^2 - C_1 \tag{66}$$

in the $e-1$ unknowns a_j. This system was seen to be equivalent to

$$p = R(1, n, \beta)R(1, n, \beta^{-1}), \qquad R(1, n, \beta) = \sum_{i=1}^{e-1} a_i\beta^i. \tag{67}$$

But* p is a product of $e-1$ prime ideals (each of norm p). When $e<23$, each ideal is a principal ideal, since there is a single class of ideals for the field F of rational functions of the primitive eth root β of unity. Hence $p=up_1 \cdots p_{e-1}$, where u is a unit of F and p_j is a polynomial in β^j with integral coefficients independent of j. Let $f(\beta)$ be a product† of $\frac{1}{2}(e-1)$ of the p_j such that $f(\beta^{-1})$ is the product of the remaining p's. Since $p_1 \cdots p_{e-1}$ is a symmetric function with integral coefficients of the roots of $\beta^{e-1}+ \cdots +\beta+1=0$, it is an integer I. By $p=uI$, $u=\pm 1$. Thus $\pm p=f(\beta)f(\beta^{-1})$. The lower sign is excluded by §4. Thus $u=1$ and

$$p = vf(\beta)\cdot v^{-1}f(\beta^{-1}), \quad v = v(\beta), \quad v^{-1} = v(\beta^{-1}).$$

The unit v is the product‡ of an integral power of β by a polynomial $P(\beta+\beta^{-1})$. Hence $P^2=1$ and $v=\pm\beta^k$. Thus

$$p = \pm F(\beta)\cdot \pm F(\beta^{-1}), \qquad F(\beta) = \beta^k f(\beta) = \sum_{j=1}^{e-1} a_j\beta^j. \tag{68}$$

Write $\beta F=\sum A_i\beta^i$, summed for $i=1, \cdots, e-1$. Then

$$A_1 = -a_{e-1}, \qquad A_{i+1} = a_i - a_{e-1} \qquad (i = 1, \cdots, e-2), \tag{69}$$

$$\sum A_j \equiv \sum a_j, \quad \sum jA_j \equiv \sum ja_j + \sum a_j \pmod{e}, \tag{70}$$

where j takes the values $1, \cdots, e-1$. Hence if $\beta^n F=\sum N_i\beta^i$,

$$\sum N_j \equiv \sum a_j, \quad \sum jN_j \equiv \sum ja_j + n\sum a_j \pmod{e}. \tag{71}$$

But $\sum a_j \not\equiv 0 \pmod{e}$ by Theorem 3 or (17). Hence in (68) there is a single value of β^k for which $\sum ja_j \equiv 0 \pmod{e}$, and hence, by (23), such that (68) is a decomposition (67) available for cyclotomy.

When β is replaced by its powers in turn, the a_i undergo the substitutions of a cyclic group of order $e-1$. These substitutions leave unaltered (23_1), and the system (23).

THEOREM 8. *If $e=5$, the eight solutions of* (66) *for which $\sum ja_j\equiv 0 \pmod 5$ are all derived from one by the powers of $(a_1a_2a_4a_3)$ and changing all signs.*

---

* Kummer. See Hilbert's report, Jahresbericht der Deutschen Mathematiker-Vereinigung, vol. 4 (1894), p. 328.

† When $e=5, f(\beta)$ is p_1p_2, p_3p_4, p_1p_3 or p_2p_4.

‡ Kummer. See Hilbert's report, p. 336.

14. **Case** $e=7$. The only factorizations $p=f(\beta)f(\beta^{-1})$ are

$$(72) \qquad p_1p_2p_3\cdot p_6p_5p_4,\quad p_1p_3p_5\cdot p_6p_4p_2,\quad p_1p_4p_5\cdot p_6p_3p_2,$$

$$(73) \qquad p_1p_2p_4\cdot p_6p_5p_3.$$

Since (73) is unaltered when β is replaced by β^2, it corresponds to (67) with $n=2$. This is true by §15 or directly by

$$(74) \quad R(1, 2) = r + s(\beta + \beta^2 + \beta^4) + t(\beta^3 + \beta^5 + \beta^6),$$

$$r = (0, 0) + 3(1, 3) + 3(1, 5),$$

$$s = (0, 1) + (0, 2) + (0, 4) + (1, 2) + (1, 4) + (1, 5) + (2, 4),$$

$$t = (0, 3) + (0, 5) + (0, 6) + (1, 2) + (1, 3) + (1, 4) + (2, 4),$$

$$(75) \qquad 4p = (s + t - 2r)^2 + 7(s - t)^2.$$

The replacement of β by β^3 induces $S=(a_1a_5a_4a_6a_2a_3)$. The same replacement carries (72_1) to (72_3) and the latter to (72_2).

Theorem 9. *If $e=7$, all solutions of* (66) *having* $\sum ja_j\equiv 0 \pmod 7$ *are of two types. For one type, the six a_i are equal in sets of three and correspond to* $R(1, 2)$. *The twelve solutions of the other type correspond to* $R(1, 1)$ *and are all derived from one by the powers of S and changing all signs.*

By (5) and (74), we find that (43) is equivalent to

$$(76) \qquad 3\sum_1^6 a_{i1} - 14r = -p - 16 - 49(0, 0), \qquad R(1, 1) = \sum_1^6 a_{i1}\beta^i,$$

since $3(s+t)=p-2-r$. Hence the first square in the decomposition (75) determines r apart from sign. The sign is fixed by the fact that (76) must yield an integer for (0, 0). This is simpler than using

$$(77) \qquad R(1, 2) = R(1, 1)R(1, 1, \beta^2)/R(1, 1, \beta^3).$$

15. Kummer* proved that, if e is a prime,

$$(78) \qquad R(1, n, \beta) = \pm\, \beta^s \prod p(\beta^{m_h}),$$

where the product extends over the $\frac{1}{2}(e-1)$ positive integers $h<e$ such that $h+[nh]>e$. Here $[x]$ denotes the least positive residue of x, and $hm_h\equiv 1 \pmod e$. Also $p(\alpha)$ is a prime ideal factor of p, and may be replaced by a polynomial in α if $e<23$.

For brevity write m for $p(\beta^m)$. Replacing β by β^{-1}, we get $e-m$.

16. **Case** $e=11$. By (78), $R(1, 1)$ and $R(1, 2)$ are the products of units $\pm\beta^j$ by the products of

* Journal für Mathematik, vol. 35 (1847), pp. 361–63; Journal de Mathématiques, vol. 12 (1847), pp. 185–212, where he gave a $p(\alpha)$ for $e=5, 7, 11, 13, 17, 19$, for all primes $p<1000$, $p\equiv 1 \pmod e$.

$$2\ 5\ 7\ 8\ 10, \qquad 3\ 5\ 7\ 9\ 10, \tag{79}$$

respectively. Replacing β by β^{-1} (i.e., subtracting each symbol from 11) in (79_1), we get the complementary set 9 6 4 3 1; the symbols in the two sets are together $1, \cdots, 10$ rearranged. In all there are 16 pairs of complementary sets. We list the sets containing 1.

I. 1 2 3 4 5	VI. 1 2 4 6 8	XI. 1 3 5 7 9
II. 1 2 3 4 6	VII. 1 2 5 7 8	XII. 1 3 6 7 9
III. 1 2 3 5 7	VIII. 1 2 6 7 8	XIII. 1 4 5 8 9
IV. 1 2 3 6 7	IX. 1 3 4 5 9	XIV. 1 4 6 8 9
V. 1 2 4 5 8	X. 1 3 4 6 9	XV. 1 5 7 8 9
		XVI. 1 6 7 8 9

Let I denote also the pair of complementary sets containing I. Likewise for II, etc. The replacement of β by β^2 permutes the pairs as follows:

$$(\mathrm{I}, \mathrm{XI}, \mathrm{XIII}, \mathrm{X}, \mathrm{VIII})\ (\mathrm{II}, \mathrm{VI}, \mathrm{V}, \mathrm{XII}, \mathrm{IV})\ (\mathrm{III}, \mathrm{XV}, \mathrm{XIV}, \mathrm{VII}, \mathrm{XVI}), \tag{80}$$

while the sets of the pair IX are interchanged. This replacement of β by β^2 induces the substitution

$$s = (a_1a_6a_3a_7a_9a_{10}a_5a_8a_4a_2) \tag{81}$$

on the coefficients of $a_1\beta + \cdots + a_{10}\beta^{10}$. Hence all solutions of (66) are obtained by applying powers of s to the solutions obtained from I, II, III, IX. We seek properties of solutions a_i which will discard the last two cases and hence retain only solutions from I and II, or if we prefer, X and VI, which occur in the cycles of (80) having I and II. Note that X and VI are complementary to (79) and hence correspond to $R(1, 1)$ and $R(1, 2)$.

The set IX is unaltered by $(\beta\beta^3\beta^9\beta^5\beta^4)$, whence the resulting ten a_i are equal in sets of five. Such immediately detected solutions of (66) yield neither $R(1, 1)$ nor $R(1, 2)$.

By (54) with $m=n=1$, $t=14$, and (9), (11), we get

$$R(1, 1, \beta) = R(1, 2, \beta^6)R(1, 2, \beta^7)/R(1, 2, \beta^3), \tag{82}$$

which is quickly verified by (8). The denominator is equal to $p/R(1, 2, \beta^8)$. Theoretically we could use (82) to show that the solution a_i obtained from III yields neither $R(1, 2)$ nor $R(1, 1)$. Practically it is simpler to verify that we do not then obtain an integral value for $(0, 0)$ by

$$3\sum a_{i1} + 6\sum a_{i2} = p - 32 - 121(0, 0), \tag{83}$$

to which (41) reduces when $e=11$ by Theorem 4.

We employ the $p(\alpha)$ in Kummer's table cited in the last foot-note. We

choose n to make the final sum in (71) a multiple of $e=11$. Then the new $\sum ja_j\equiv 0 \pmod{11}$. Finally we change all signs (if necessary) to make $\sum a_i\equiv -1$ as in (22).

$$p = 683,\quad p(\beta) = 2+\beta$$

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	$\sum a_i$
$-\beta^6$ I	12	12	16	16	22	−6	10	14	6	7	109
$-\beta^2$ II	2	−24	−6	−6	−10	−10	−1	−12	−4	4	−67
$-\beta^5$ III	−11	−6	2	−2	−24	−14	−8	−14	0	−12	−89

Then (83) holds modulo 121 if and only if

$$\left(\sum a_{i1}, \sum a_{i2}\right) = (109, -67) \text{ or } (-89, -89).$$

The second case is excluded since the two products (79) are distinct. Hence the first and second rows of our tablette give the coefficients of $R(1, 1)$ and $R(1, 2)$ respectively.* Also, $(0, 0)=6$.

$$p=991,\quad p(\beta)=2+\beta+\beta^3$$

$-\beta^7$ I	8	8	4	29	10	18	6	18	26	4	$\sum = 131$
$-\beta^6$ II	−6	−14	18	−10	−13	−16	−12	−8	−4	−2	$\sum = -67$
$-\beta^4$ III	10	4	−10	−12	4	−4	−19	2	−18	−2	$\sum = -45$

Here (83) holds only when the first row gives $R(1, 1)$ and the second $R(1, 2)$. Thus $(0, 0)=8$.

$$p = 199,\quad p(\beta) = 1+\beta-\beta^2$$

$-\beta^4$ I	2	2	−4	4	6	8	4	−5	6	−2	$\sum = 21$
$-\beta^5$ II	2	−2	−8	−10	4	−6	0	0	−2	−1	$\sum = -23$
$-\beta^7$ III	6	6	7	8	16	2	2	6	4	8	$\sum = 65$

The first row gives $R(1, 1)$, the second $R(1, 2)$, the third is excluded by (83).

$p=23$, $p(\beta)=1+\beta+\beta^9$. Then $-\beta^5$ I, β^9 II, $-\beta^6$ III have $\sum a_i=21, -12, -1$. Here (83) holds only when the sums are 21 and -12 or both -1, the last contrary to (79).

We have treated the four p's <1000 for which $p(\beta)$ has fewer than four terms.

17. **Case** $e=13$. By (78), $R(1, 1)$, $R(1, 2)$ and $R(1, 3)$ are the products of units $\pm\beta^s$ by the products of

* Apart from a power of (81), depending on the root β chosen. Likewise for the similar later statements.

(84) $\qquad 2, 3, 4, 5, 6, 12;\ 3, 4, 6, 8, 11, 12;\ 2, 4, 5, 6, 10, 12,$

respectively. Since $R(1, 3)$ is unaltered when β is replaced by β^3, its 12 coefficients a_i are equal in sets of three. No other factorization of p has this property.

The replacement of β by β^2 gives rise to a cyclic substitution S on the twelve powers of β. Applying S, S^3, S to the sets complementary (§16) to (84), we get

(85) $\qquad A = 1, 2, 3, 5, 7, 9;\ B = 1, 2, 3, 4, 7, 8;\ C = 1, 2, 3, 5, 6, 9.$

There are exactly 32 decompositions of p into two complementary products of six factors. Of them, (84_3) and C are permuted by S. The others form five cycles of six, those of a cycle being permuted by S. It suffices to know one entry from each cycle. We may take them to be A, B and

(86) $\qquad D = 1, 2, 3, 4, 5, 6;\ E = 1, 2, 3, 4, 5, 7;\ F = 1, 3, 4, 5, 6, 11.$

The choice $A, \cdots, F$ facilitates forming and checking the products. Each product of six is multiplied by a unit $\pm\beta^s$ uniquely determined (§13) by

$$\sum a_j \equiv -1, \qquad \sum j a_j \equiv 0 \pmod{13}.$$

Theoretically it would be possible to prove by means of

$$(87)\quad R(1, 1) = \frac{R(1, 2, \beta^4)R(1, 2, \beta^7)}{R(1, 2, \beta^2)}, \qquad R(1, 3) = \frac{R(1, 2, \beta^7)R(1, 2, \beta^8)}{R(1, 2, \beta^2)}$$

that $R(1, 1)$, $R(1, 2)$, $R(1, 3)$ correspond to A, B, C, respectively, while (86) are excluded. Practically we make use of

$$(88)\qquad 3\sum a_{i1} + 6\sum a_{i2} + 2\sum a_{i3} = p - 38 - 169(0, 0),$$

to which (41) reduces when $e = 13$ by Theorem 4.

$$p = 79,\ p(\beta) = 1 - \beta + \beta^{10}$$

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}	$\sum a_i$
$-\beta^4 A$	−6	2	0	0	−2	−4	−2	−2	0	−4	−6	−3	−27
$-\beta^9 B$	0	−3	−5	−1	−6	−3	−1	−3	3	−1	−2	−5	−27
$-C$	−1	−5	−1	−4	−5	−5	1	1	−1	−4	1	−4	−27
$-\beta^6 D$	2	4	4	2	−3	0	3	0	−2	3	−3	2	12
$-\beta^{10} E$	4	3	8	3	4	4	4	8	8	4	7	7	64
$-\beta^3 F$	−2	−4	−5	1	−2	4	−1	−2	1	0	−4	0	−14

Since $\sum a_{i3} = -27$, (88) holds modulo 169 only when the three sums are all -27 or are 64, 12, -27, respectively. But the latter case gives $(0, 0) = -1$ and is excluded. Hence the first three rows of the table give the coefficients of $R(1, 1)$, $R(1, 2)$, $R(1, 3)$, while $(0, 0) = 2$.

$$p = 521, \quad p(\beta) = 1 + \beta - \beta^{12}$$

$-\beta^{11}A$	-2	4	-4	0	12	4	4	-6	16	8	7	8	51
$-\beta^{2}B$	18	11	18	4	4	0	6	10	6	2	16	8	103
$-C$	-15	-13	-15	-3	-13	-13	-17	-17	-15	-3	-17	-3	-144
$-\beta^{10}D$	4	-4	-16	0	-4	-8	-8	-10	-4	-19	-4	-6	-79
$-\beta^{8}E$	12	-12	-4	-6	-4	6	2	-5	-2	6	4	2	-1
$-\beta^{5}F$	-18	-10	-10	0	-5	-8	2	-6	-18	-12	-8	-12	-105

Since $\sum a_{i3} = -144$, (88) holds modulo 169 only when the first two sums are 51 and 103, or -79 and -1. The latter case gives $(0, 0) = +6$ and is to be excluded by other means such as (87).

$$p = 131, \quad p(\beta) = 1 - \beta + \beta^{11}$$

$-\beta^{3}A$	2	4	4	0	6	1	0	2	-6	2	4	6	25
$-\beta^{10}B$	5	-5	4	2	1	0	3	4	4	0	-2	-4	12
$-C$	-3	-5	-3	3	-5	-5	-4	-4	-3	3	-4	3	-27
$-\beta^{11}D$	-7	-3	1	-5	-3	-6	3	0	2	-1	-5	-3	-27
$-\beta E$	2	3	1	-2	-1	-2	0	2	2	2	9	-4	12
$-\beta^{12}F$	3	0	2	6	6	-6	1	-2	0	2	-1	1	12

Then (88) holds modulo 169 if the third sum is -27 only when the first sum is 25 and the second is 12.

University of Chicago,
Chicago, Ill.

ON THE EXPANSION OF THE REMAINDER IN THE NEWTON-COTES FORMULA*

BY
J. V. USPENSKY

1. In Newton's method for approximate evaluation of definite integrals the interval of integration, say (0, 1), is divided into a certain number n of equal parts and the integral of a given function $f(x)$ is assumed to be approximately equal to the integral of the interpolation polynomial of degree n which at the points of subdivision has the same values as $f(x)$. The resulting approximate formula

$$\int_0^1 f(x)dx = A_0 f(0) + A_1 f\left(\frac{1}{n}\right) + \cdots + A_n f(1) \tag{1}$$

is known as the Newton-Cotes quadrature formula. The coefficients A_0, $A_1, \cdots, A_n$ depend on the number of divisions n and their values have been computed by Cotes for $n \leqq 10$. In the following we shall call them "Cotes coefficients."

Formula (1) is exact for an arbitrary polynomial $f(x)$ of degree not exceeding n. However, since for an *even* n

$$\int_0^1 x\left(x - \frac{1}{n}\right)\left(x - \frac{2}{n}\right) \cdots (x - 1)dx = 0,$$

formula (1) will be exact even for polynomials of degree $n+1$ if n is even. Strange as it may seem, the expression for the remainder in the Newton-Cotes formula was unknown till quite recently. It was only in 1922 and 1924 that J. F. Steffenson succeeded in giving a genuine expression of that remainder first for an even and then for an odd n.† In this paper we shall show that the remainder in the Newton-Cotes formula possesses an expansion in all respects quite similar to the classical Euler-Maclaurin expansion, which fact is interesting in itself and may be very useful in practice. The method by which this result is obtained is similar to that employed in our paper *On an expansion of the remainder in the Gaussian quadrature formula*,‡ but besides

* Presented to the Society, September 7, 1934; received by the editors August 9, 1934.

† The detailed derivation of the remainder in the Newton-Cotes formula can be found in an excellent book by J. F. Steffenson, *Interpolation*, Baltimore, Williams and Wilkins, 1927.

‡ Bulletin of the American Mathematical Society, vol. 40 (1934).

simple means used in that paper it requires an elaborate study of certain properties of Cotes coefficients.*

2. In a paper *Sur les valeurs asymptotiques des coefficients de Cotes*† I have shown that for large n and uniformly in k $(1\leqq k\leqq n-1)$

$$A_k \sim \frac{C_n^k}{n(\log n)^2}\left[\frac{(-1)^{k-1}}{k}+\frac{(-1)^{n-k-1}}{n-k}\right],$$

while

$$A_0 = A_n \sim \frac{1}{n\log n}\cdot$$

These formulas show that for sufficiently large n

(2) $$(-1)^{k-1}A_k > 0,\ A_0 > 0,$$

if $1\leqq k\leqq \frac{1}{2}n$; as to the signs of the remaining coefficients, they result from the elementary relation

$$A_{n-k} = A_k.$$

Now it is very important to determine the least value of n for and after which inequalities (2) hold. It will be shown that $A_0>0$ always and that $(-1)^{k-1}A_k>0$ $(1\leqq k\leqq \frac{1}{2}n)$ for even $n\geqq 8$ and odd $n\geqq 11$.

3. From our article quoted above‡ we take the following expression of $A_k(1\leqq k\leqq n-1)$:

$$A_k = \frac{1}{n}+\frac{n+1}{n}C_n^k\,(-1)^{k-1}\left[\int_0^1 \xi^k(1-\xi)^{n-k}d\xi\int_{-\infty}^{\log((1-\xi)/\xi)}\frac{e^{kx}}{\pi^2+x^2}dx\right.$$
$$\left.+(-1)^n\int_0^1\xi^{n-k}(1-\xi)^k d\xi\int_{-\infty}^{\log((1-\xi)/\xi)}\frac{e^{(n-k)x}}{\pi^2+x^2}dx\right].$$

Introducing here instead of x a new variable t defined by

$$\frac{\xi}{1-\xi}e^x = t,$$

we can present A_k in the form

* I must express my thanks to Professor G. Pólya for some very helpful suggestions in connection with this investigation.

† Bulletin of the American Mathematical Society, vol. 31 (1925), pp. 145–156. See also G. Pólya, *Ueber die Konvergenz von Quadraturverfahren*, Mathematische Zeitschrift, vol. 37(1933).

‡ Attention should be called to the fact that many formulas on p. 147 of that article are marred by typographical errors.

$$(3) \qquad A_k = \frac{1}{n} + \frac{n+1}{n}(-1)^{k-1}C_n^k \int_0^1 \int_0^1 \frac{\phi(t)(1-\xi)^n d\xi dt}{\pi^2 + \left(\log \frac{1-\xi}{\xi} t\right)^2},$$

$$\phi(t) = t^{k-1} + (-1)^n t^{n-k-1},$$

whence

$$(-1)^{k-1} n A_k = (-1)^{k-1} + (n+1)C_n^k \int_0^1 \int_0^1 \frac{\phi(t)(1-\xi)^n d\xi dt}{\pi^2 + \left(\log \frac{1-\xi}{\xi} t\right)^2}$$

and

$$(-1)^{k-1} n A_k > (-1)^{k-1} + (n+1)C_n^k \int_0^{1/2} (1-\xi)^n F(\xi) d\xi,$$

$$F(\xi) = \int_{\xi/(1-\xi)}^1 \frac{t^{k-1} + (-1)^n t^{n-k-1}}{\pi^2 + \left(\log \frac{1-\xi}{\xi} t\right)^2} dt.$$

From this inequality it follows immediately that for an *odd* $k \leqq n/2$ the coefficient A_k is positive.

Integrating by parts we find

$$\int_{\xi/(1-\xi)}^1 \frac{t^{k-1} + (-1)^n t^{n-k-1}}{\pi^2 + \left(\log \frac{1-\xi}{\xi} t\right)^2} dt > \frac{\frac{1}{k} + \frac{(-1)^n}{n-k}}{\pi^2 + \left(\log \frac{1-\xi}{\xi}\right)^2}$$

$$- \frac{1}{\pi^2}\left[\frac{1}{k}\left(\frac{\xi}{1-\xi}\right)^k + \frac{(-1)^n}{n-k}\left(\frac{\xi}{1-\xi}\right)^{n-k}\right],$$

and hence

$$(n+1)C_n^k \int_0^{1/2} (1-\xi)^n d\xi \int_{\xi/(1-\xi)}^1 \frac{t^{k-1} + (-1)^n t^{n-k-1}}{\pi^2 + \left(\log \frac{1-\xi}{\xi} t\right)^2} dt$$

$$> (n+1)C_n^k \left(\frac{1}{k} + \frac{(-1)^n}{n-k}\right) \int_0^{1/2} \frac{(1-\xi)^n d\xi}{\pi^2 + \left(\log \frac{1-\xi}{\xi}\right)^2} - \frac{1}{k\pi^2}$$

because

$$\frac{(n+1)C_n^k}{k} \int_0^{1/2} \xi^k (1-\xi)^{n-k} d\xi + (-1)^n \frac{(n+1)C_n^k}{n-k} \int_{1/2}^1 \xi^k (1-\xi)^{n-k} d\xi$$

$$< \frac{(n+1)C_n^k}{k} \int_0^1 \xi^k (1-\xi)^{n-k} d\xi = \frac{1}{k}.$$

Consequently for an even $k \leqq n/2$

$$(4)\quad -nA_k > (n+1)C_n^k\left(\frac{1}{k}+\frac{(-1)^n}{n-k}\right)\int_0^{1/2}\frac{(1-\xi)^n d\xi}{\pi^2+\left(\log\frac{1-\xi}{\xi}\right)^2}-1-\frac{1}{k\pi^2}.$$

4. In order to draw the desired conclusions from this inequality we must find a sufficiently precise and tractable lower limit of the integral

$$K_n = \int_0^{1/2}\frac{(1-\xi)^n}{\pi^2+\left(\log\frac{1-\xi}{\xi}\right)^2}d\xi.$$

Consider a more general integral

$$K_n(\alpha) = \int_0^{1/2}\frac{\xi^{-\alpha}(1-\xi)^{n+\alpha}}{\pi^2+\left(\log\frac{1-\xi}{\xi}\right)^2}d\xi, \qquad -n < \alpha \leqq 0,$$

which for $\alpha = 0$ reduces to K_n. Since

$$\frac{d^2K_n(\alpha)}{d\alpha^2}+\pi^2K_n(\alpha) = \int_0^{1/2}(1-x)^{n+\alpha}x^{-\alpha}dx = \phi(\alpha),$$

we obtain, by integrating this equation,

$$K_n(\alpha) = K_n(-2)\cos\pi\alpha+\frac{K_n'(-2)}{\pi}\sin\pi\alpha+\frac{1}{\pi}\int_{-2}^{\alpha}\phi(x)\sin\pi(\alpha-x)dx.$$

Taking here $\alpha = 0$ we get

$$K_n = K_n(-2)+\frac{1}{\pi}\int_0^2\phi(-x)\sin\pi x dx.$$

But

$$\phi(-x) = \int_0^{1/2}\xi^x(1-\xi)^{n-x}d\xi = \frac{\Gamma(1+x)\Gamma(n+1-x)}{\Gamma(n+2)}-\frac{\vartheta}{2^n n},$$
$$0<\vartheta<1 \text{ if } n \geqq 5,$$

and $K_n(-2) > 0$; consequently

$$K_n > \frac{1}{\pi}\int_0^2\omega(x)\sin\pi x dx-\frac{1}{2^{n-1}n\pi^2},$$

where

$$\omega(x) = \frac{\Gamma(1+x)\Gamma(n+1-x)}{\Gamma(n+2)}.$$

As

$$\omega(x) - \omega(x+1) = \frac{n-1-2x}{n-x}\omega(x),$$

we shall have a fortiori

$$K_n > \frac{n-3}{(n-1)\pi}\int_0^1 \omega(x)\sin \pi x dx - \frac{1}{2^{n-1}n\pi^2}\cdot$$

On the other hand $\Gamma(1+x) > e^{-Cx}$, where C is Euler's constant,

$$\frac{\Gamma(n+1-x)}{\Gamma(n+1)} > e^{-x\log(n+1)},$$

and

$$\frac{1}{\pi}\int_0^1 e^{-x[C+\log(n+1)]}\sin \pi x dx = \frac{1}{\pi^2 + \{C+\log(n+1)\}^2}\left(1+\frac{e^{-C}}{n+1}\right).$$

Hence, finally,

$$(n+1)K_n > \frac{n-3}{n-1}\,\frac{1}{\pi^2+\{C+\log(n+1)\}^2}\left(1+\frac{e^{-C}}{n+1}\right) - \frac{n+1}{2^{n-1}n\pi^2},$$

and, taking into account inequality (4),

$$-nA_k > C_n^k\left(\frac{1}{k}+\frac{(-1)^n}{n-k}\right)\left[\frac{n-3}{n-1}\,\frac{1}{\pi^2+\{C+\log(n+1)\}^2}\left(1+\frac{e^{-C}}{n+1}\right)\right.$$
$$\left. - \frac{n+1}{2^{n-1}n\pi^2}\right] - 1 - \frac{1}{k\pi^2}\cdot$$

To prove that, for certain n and even $k \leqq n/2$, coefficients A_k are negative, it suffices to verify the inequality

$$(5)\qquad \begin{aligned} &C_n^k\left(\frac{1}{k}+\frac{(-1)^n}{n-k}\right)\left[\frac{n-3}{n-1}\,\frac{1}{\pi^2+\{C+\log(n+1)\}^2}\left(1+\frac{e^{-C}}{n+1}\right)\right.\\ &\qquad\qquad \left. - \frac{n+1}{2^{n-1}n\pi^2}\right] - 1 - \frac{1}{k\pi^2} > 0.\end{aligned}$$

5. If n is even the inequality (5) will be verified for all even values of k if it is verified for $k=2$; also, being verified for some even n, it will be verified for greater values of n. Now, taking $n=12$ and $k=2$, inequality (5) is verified; hence for an even $n \geqq 12$, A_k with even subscripts $\leqq n/2$ are all negative. The same being true for $n=8$ and $n=10$ we may be assured that

$$(-1)^{k-1}A_k > 0, \qquad 1 \leqq k \leqq \frac{n}{2},$$

for all even $n \geqq 8$. If n is odd $\geqq 13$ it suffices again to verify inequality (5) for $n = 13$, $k = 2$; by direct computation we find that it is satisfied. Hence

$$(-1)^{k-1}A_k > 0, \qquad 1 \leqq k \leqq \frac{n-1}{2},$$

for all odd $n \geqq 13$. It remains to be seen whether this is true for $n = 11$. But by direct computation I have found the following values of Cotes coefficients for $n = 11$:

$$A_0 = A_{11} = \frac{2171465}{24 \times 10!}; \quad A_1 = A_{10} = \frac{13486539}{24 \times 10!}; \quad A_2 = A_9 = -\frac{3237113}{24 \times 10!};$$
$$A_3 = A_8 = \frac{25226685}{24 \times 10!}; \quad A_4 = A_7 = -\frac{9595542}{24 \times 10!}; \quad A_5 = A_6 = \frac{15493566}{24 \times 10!}.$$

Hence the above property of Cotes coefficients holds for all odd $n \geqq 11$, but it no longer holds for $n = 9$. It remains to show that A_0 is always positive. The sign of A_0 is the same as the sign of the integral

$$I_n = \int_0^n (1-x)(2-x) \cdots (n-x)dx,$$

which, by setting

$$\Phi(x) = \int_0^x t(1-t)(2-t) \cdots (n-t)dt,$$

can be presented thus:

$$I_n = \frac{\Phi(n)}{n} + \int_0^n \frac{\Phi(x)}{x^2}dx.$$

But it is known that $\Phi(x) > 0$ in the interval $0 < x < n$, hence $I_n > 0$ and $A_0 > 0$.

6. For $n \equiv 3 \pmod 4$ it is important to show that

$$\sum_{k=0}^{(n-1)/2} kA_k > \frac{n}{8}$$

for $n \geqq 11$. To present the left hand member in a convenient form we may use formula (3). Thus we get

$$\sum_{k=0}^{(n-1)/2} kA_k = \frac{n^2-1}{8n} + \frac{n+1}{n}\int_0^1 \int_0^1 \frac{(1-\xi)^n \phi(t)}{\pi^2 + \left(\log \dfrac{1-\xi}{\xi} t\right)^2} d\xi dt,$$

where

$$\phi(t) = \sum_{k=0}^{(n-1)/2} (-1)^{k-1} C_n{}^k k (t^{k-1} - t^{n-k-1}).$$

By an elementary transformation this expression becomes

$$\phi(t) = n(1-t)\psi(t),$$

$\psi(t)$ denoting a reciprocal polynomial

$$\psi(t) = 1 - C_{n-2}^{1} t + \cdots - C_{n-2}^{(n-5)/2} t^{(n-5)/2} \\ + C_{n-2}^{(n-3)/2} t^{(n-3)/2} - C_{n-2}^{(n-5)/2} t^{(n-1)/2} + \cdots + t^{n-3}.$$

Using the following known identity,

$$t^{n-2} - C_{n-2}^{1} t^{n-3} + \cdots - C_{n-2}^{(n-5)/2} t^{(n+1)/2} \\ = -\frac{n-3}{2} C_{n-2}^{(n-3)/2} (1-t)^{n-2} \int_0^t \frac{y^{(n-1)/2}dy}{(1-y)^{n-1}},$$

$\psi(t)$ can be presented thus:

$$\psi(t) = -\frac{n-3}{2} C_{n-2}^{(n-3)/2} (1-t)^{n-2} (1+t^{-1}) \int_0^t \frac{y^{(n-1)/2}dy}{(1-y)^{n-1}} \\ + (1-t)^{n-2} + C_{n-2}^{(n-3)/2} t^{(n-1)/2},$$

whence

$$\frac{d}{dt} \frac{t\psi(t)}{(1+t)(1-t)^{n-2}} = \frac{2t^{(n-1)/2}}{(1-t)^{n-1}(1+t)^2} C_{n-2}^{(n-3)/2} + \frac{1}{(1+t)^2} > 0; \ t > 0.$$

It follows that $\psi(t) > 0$ for all real values of t and the same is true of $\phi(t)$ if $t < 1$.

This analysis shows incidentally that the algebraic equation

$$1 - C_{n-2}^{1} t + \cdots - C_{n-2}^{(n-5)/2} t^{(n-5)/2} + C_{n-2}^{(n-3)/2} t^{(n-3)/2} \\ - C_{n-2}^{(n-5)/2} t^{(n-1)/2} + \cdots + t^{n-3} = 0$$

has only imaginary roots if $n \equiv 3 \pmod 4$; but if $n \equiv 1 \pmod 4$ it has exactly two real roots as can be shown by similar considerations. Since

$$\phi(t) = n(1-t)^{n-1} + nC_{n-2}^{(n-3)/2} t^{(n-1)/2} (1-t) \\ - \frac{n(n-3)}{2} C_{n-2}^{(n-3)/2} (1-t)^{n-1} (1+t^{-1}) \int_0^t \frac{y^{(n-1)/2}dy}{(1-y)^{n-1}}$$

remains positive in the interval $0<t<1$ it follows that

$$
(6)\qquad \sum_{k=0}^{(n-1)/2} kA_k > \frac{n^2-1}{8n} + \frac{n+1}{n}\int_0^{1/2}(1-\xi)^n G(\xi)d\xi; \qquad G(\xi) = \int_{\xi/(1-\xi)}^{1} \frac{\phi(t)}{\pi^2 + \left(\log\frac{1-\xi}{\xi}t\right)^2}dt.
$$

But integrating by parts we find

$$
(7)\qquad \int_{\xi/(1-\xi)}^{1} \frac{\phi(t)dt}{\pi^2 + \left(\log\frac{1-\xi}{\xi}t\right)^2} > \frac{\Phi(1)}{\pi^2 + \left(\log\frac{1-\xi}{\xi}\right)^2} - \frac{\Phi\left(\frac{\xi}{1-\xi}\right)}{\pi^2},
$$

where

$$
\Phi(x) = \int_0^x \phi(t)dt.
$$

But since

$$
\phi(t) > n(1-t)^{n-1} + nC_{n-2}^{(n-3)/2}t^{(n-1)/2}(1-t) - \frac{n(n-3)}{2}C_{n-2}^{(n-3)/2}(1-t)^{n-1}\int_0^t \frac{y^{(n-1)/2}+y^{(n-3)/2}}{(1-y)^{n-1}}dy,
$$

we find, after performing simple integrations,

$$
(8)\qquad \Phi(1) > 1 + \frac{4C_{n-2}^{(n-3)/2}}{(n+1)(n-1)}\cdot
$$

On the other hand

$$
\phi(t) < nC_{n-2}^{(n-3)/2}\left(1-\frac{n-3}{n-2}\right)t^{(n-1)/2}(1-t) + n(1-t)^{n-1} + \frac{n(n-1)}{2(n-2)}C_{n-2}^{(n-3)/2}t^{(n-3)/2}(1-t)^2,
$$

whence

$$
\Phi(t) < \frac{2nC_{n-2}^{(n-3)/2}}{(n+1)(n-2)}t^{(n+1)/2} + 1 - (1-t)^n + \frac{n(n-1)}{n-2}C_{n-2}^{(n-3)/2}\left\{\frac{t^{(n-1)/2}}{n-1} - \frac{2t^{(n+1)/2}}{n+1} + \frac{t^{(n+3)/2}}{n+3}\right\},
$$

and after simple calculations,

$$\int_0^{1/2} \Phi\left(\frac{\xi}{1-\xi}\right)(1-\xi)^n d\xi < \frac{1}{2(n-2)}. \tag{9}$$

Inequalities (6), (7), (8), (9) combined give

$$\sum_{k=0}^{(n-1)/2} kA_k - \frac{n}{8} > \left[\frac{n+1}{n} + \frac{4C_{n-2}^{(n-3)/2}}{n(n-1)}\right] \int_0^{1/2} \frac{(1-\xi)^n d\xi}{\pi^2 + \left(\log \dfrac{1-\xi}{\xi}\right)^2} - \frac{1}{8n} - \frac{n+1}{2\pi^2 n(n-2)}. \tag{10}$$

Now using the lower limit of the integral

$$\int_0^{1/2} \frac{(1-\xi)^n d\xi}{\pi^2 + \left(\log \dfrac{1-\xi}{\xi}\right)^2}$$

given in §5 we find that the right hand side of the inequality (10) is positive for $n=15$, and so

$$\sum_{k=0}^{(n-1)/2} kA_k > \frac{n}{8}$$

for $n \geqq 15$. But this statement is true already for $n=11$ as can be ascertained directly by using values of Cotes coefficients for $n=11$ given in §5.

7. To derive the expansion of the remainder R_n in the Newton-Cotes formula we take for starting point the well known identity valid for $0 \leqq \theta \leqq 1$

$$f(\theta) = \int_0^1 f(x)dx + \sum_{s=1}^{l} \frac{B_s(\theta)}{s!}\left\{f^{(s-1)}(1) - f^{(s-1)}(0)\right\} - \int_0^1 \frac{\overline{B}_l(\theta - t)}{l!} f^{(l)}(t)dt,$$

in which $B_n(x)$ denotes the Bernoullian polynomial of order n (defined as in Nörlund's *Differenzenrechnung*) and $\overline{B}_n(x)$ is a periodic function:

$$\overline{B}_n(x) = B_n(x) \text{ for } 0 \leqq x < 1,$$

$$\overline{B}_n(x+1) = \overline{B}_n(x) \text{ for all } x.$$

Take $l=2\nu=n+1$ or $n+2$ according as n is odd or even. Set $\theta=0,\ 1/n,\ 2/n, \cdots, 1$, multiply the resulting equations by the corresponding Cotes coefficients and take the sum; the result will be

$$\sum_{i=0}^{n} A_i f\left(\frac{i}{n}\right) = \int_0^1 f(x)dx + \frac{\sum_{i=0}^{n} A_i B_{2\nu}\left(\frac{i}{n}\right)}{(2\nu)!}\left\{f^{(2\nu-1)}(1) - f^{(2\nu-1)}(0)\right\}$$
$$(11)\qquad - \int_0^1 \frac{f^{(2\nu)}(t)}{(2\nu)!}\sum_{i=0}^{n} A_i \overline{B}_{2\nu}\left(\frac{i}{n} - t\right)dt,$$

since for $s=1, 2, \cdots, 2\nu-1$

$$\sum_{i=0}^{n} A_i B_s\left(\frac{i}{n}\right) = \int_0^1 B_s(x)dx = 0.$$

Setting for brevity

$$\sum_{i=0}^{n} A_i\left\{\overline{B}_p\left(\frac{i}{n} - t\right) - B_p\left(\frac{i}{n}\right)\right\} = G_p(t)$$

it follows from (11) that

$$(12)\qquad R_n = \frac{1}{(2\nu)!}\int_0^1 G_{2\nu}(t)f^{(2\nu)}(t)dt.$$

From the definition of $G_p(t)$ we see that

$$G_p(0) = G_p(1) = 0,$$

and without any difficulty (owing to the fact that $A_{n-i}=A_i$) we can establish the following relations:

$$G_p(1 - t) = (-1)^p G_p(t), \quad \sum_{i=0}^{n} A_i B_{2s-1}\left(\frac{i}{n}\right) = 0,$$

so that $G_{2s-1}(t)$ can be written simply thus:

$$G_{2s-1}(t) = \sum_{i=0}^{n} A_i \overline{B}_{2s-1}\left(\frac{i}{n} - t\right),$$

and from this expression, because

$$\overline{B}_n'(x) = n\overline{B}_{n-1}(x),$$

the following relations can be derived:

$$(13)\qquad G_{2s}'(t) = -2sG_{2s-1}(t),$$

$$(14)\qquad G_{2s+1}''(t) = 2s(2s + 1)G_{2s-1}(t),$$

$$(15)\qquad G_{2s}''(t) = 2s(2s - 1)\left[G_{2s-2}(t) + \sum_{i=0}^{n} A_i B_{2s-2}\left(\frac{i}{n}\right)\right].$$

8. For $s \geqq \nu$ functions $G_{2s}(t)$ do not change sign in the interval $0<t<1$. To prove this fundamental property, let β_s and α_s represent, respectively, the number of times $G_{2s}(t)$ and $G_{2s-1}(t)$ change their sign when t increases from 0 to 1. Because $G_{2s}(0)=G_{2s}(1)=0$ it follows from (13) and Rolle's theorem that

$$\beta_s + 1 \leqq \alpha_s.$$

Again, using (14) and applying Rolle's theorem twice, we get

$$\alpha_s \leqq \alpha_{s-1},$$

so that for $s \geqq \nu$,

$$\beta_s + 1 \leqq \alpha_\nu.$$

Since

$$G_{2\nu-1}(1-t) = -G_{2\nu-1}(t),$$

it follows that $\alpha_\nu \geqq 1$ and it is important to prove that $\alpha_\nu = 1$.

To this end we notice first that for $0 \leqq t \leqq 1$

$$\begin{aligned} G_{2\nu-1}(t) &= \sum_{i=0}^{n} A_i \overline{B}_{2\nu-1}\left(\frac{i}{n} - t\right) \\ &= \sum_{i=0}^{n} A_i B_{2\nu-1}\left(\frac{i}{n} - t\right) + (2\nu-1)\sum_{i \leqq nt} A_i\left(\frac{i}{n} - t\right)^{2\nu-2}. \end{aligned}$$

Furthermore, $B_{2\nu-1}(x)$ being a polynomial of degree $2\nu-1$,

$$\begin{aligned} \sum_{i=0}^{n} A_i B_{2\nu-1}\left(\frac{i}{n} - t\right) &= \int_0^1 B_{2\nu-1}(x-t)dx \\ &= \frac{1}{2\nu}\{B_{2\nu}(1-t) - B_{2\nu}(-t)\} = -t^{2\nu-1}. \end{aligned}$$

Hence $G_{2\nu-1}(t)$ differs only by a constant factor from the function

$$R_0(t) = \frac{t^{2\nu-1}}{2\nu-1} - \sum_{i \leqq nt} A_i\left(\frac{i}{n} - t\right)^{2\nu-2},$$

and the number of times this function changes sign in the interval $0<t<1$ is α_ν.

We shall now prove the following Fundamental Lemma:

FUNDAMENTAL LEMMA. *The function $R_0(t)$ changes sign once and only once in the interval $0<t<1$, so that $\alpha_\nu = 1$.*

Let

$$R_k(t) = \frac{(-1)^k t^{2\nu-k-1}}{2\nu - k - 1} - \sum_{i \leqq nt} A_i \left(\frac{i}{n} - t\right)^{2\nu-k-2}$$

for $k=0, 1, 2, \cdots, 2\nu-2$. These functions differ only by constant factors from the successive derivatives of $R_0(t)$. With the exception of $R_{2\nu-2}(t)$ they are continuous, but

$$R_{2\nu-2}(t) = t - \sum_{i \leqq nt} A_i$$

has discontinuities at the points $1/n, 2/n, \cdots, (n-1)/n$. By the fundamental property of the Newton-Cotes formula,

$$R_k(1) = R_k(0) = 0.$$

Also

$$R_k(1 - t) = (-1)^{k-1} R_k(t),$$

whence it follows that for an even subscript k there is always a change of sign at $t=\frac{1}{2}$.

Let N_k in general represent the number of changes of sign of $R_k(t)$ in the interval $0<t<1$. Then by Rolle's theorem

$$N_k + 1 \leqq N_{k+1}$$

for $k=0, 1, 2, \cdots, 2\nu-3$. Hence

$$N_0 + 2\nu - 2 \leqq N_{2\nu-2}$$

and also

$$N_0 + 2\nu - 3 \leqq N_{2\nu-3}.$$

Now it is possible to assign a certain upper limit to $N_{2\nu-2}$. To this end we distinguish the following cases:

Case 1: $n=2\nu-2\geqq 8$ *and divisible by* 4. The interval $0<t<\frac{1}{2}$ may be divided into $(\nu+1)/2$ intervals:

$$0 < t < \frac{1}{n};\ \frac{1}{n} < t < \frac{3}{n};\ \frac{3}{n} < t < \frac{5}{n};\ \cdots;\ \frac{\nu-4}{n} < t < \frac{\nu-2}{n};$$

$$\frac{\nu-2}{n} < t < \frac{\nu-1}{n}.$$

In the interval $0<t<1/n$, since $A_0>0$, $R_{2\nu-2}(t)$ might change sign, but not more than once. In the last interval $(\nu-2)/n<t<(\nu-1)/n$ no changes of sign occur, since $R_{2\nu-2}(t)$ *changes* its sign at $t=\frac{1}{2}$ and $A_{\nu-1}<0$. In each of the intervals of the type $(2k-1)/n<t<(2k+1)/n$ the sign may change at most

once, since $A_{2k}<0$. Considering that $R_{2\nu-2}(t)$ may change sign at end points of the intervals, the total number of changes of sign in the interval $0<t<\frac{1}{2}$ cannot exceed $\nu-1$; in the interval $\frac{1}{2}<t<1$ there is exactly the same number of changes of sign and the sign changes at $t=\frac{1}{2}$. Hence $N_{2\nu-2}\leqq 2\nu-1$ and $N_0+2\nu-2\leqq N_{2\nu-2}$ and this necessarily implies $N_0=1$, $N_{2\nu-2}=2\nu-1$.

Case 2: $n=2\nu-2\geqq 10$ *and* $n\equiv 2 \pmod 4$. By very similar considerations we find again $N_{2\nu-2}\leqq 2\nu-1$, whence $N_0=1$, $N_{2\nu-2}=2\nu-1$.

Case 3: $n=2\nu-1\geqq 13$, ν *odd*. Divide again the interval $0<t<\frac{1}{2}$ into $(\nu+3)/2$ parts:

$$0 < t < \frac{1}{n}; \quad \frac{1}{n} < t < \frac{3}{n}; \cdots ; \frac{\nu-4}{n} < t < \frac{\nu-2}{n}; \quad \frac{\nu-2}{n} < t < \frac{\nu-1}{n};$$

$$\frac{\nu-1}{n} < t < \frac{1}{2}\cdot$$

In the first interval there might be at most one change of sign. There are no changes of sign in the last interval, since $R_{2\nu-2}(t)$ passes at $t=\frac{1}{2}$ from negative to positive values. Neither does the sign change in the interval $(\nu-2)/n<t<(\nu-1)/n$, since at the right end of this interval $R_{2\nu-2}(t)$ is negative. Altogether $R_{2\nu-2}(t)$ in the interval $0<t<\frac{1}{2}$ cannot have more than $(\nu-1)/2+(\nu-1)/2=\nu-1$ changes of sign. Hence $N_{2\nu-2}\leqq 2\nu-1$ and again $N_0=1$, $N_{2\nu-2}=2\nu-1$.

Case 4: $n=2\nu-1\geqq 11$, ν *even*. In a similar manner we find $2\nu+1$ as the upper limit of $N_{2\nu-2}$. The inequality $N_0+2\nu-2\leqq 2\nu+1$ shows only that either $N_0=1$ or $N_0=3$. To remove the last possibility we take into consideration $R_{2\nu-3}(t)$. First from the inequality $N_{2\nu-3}+1\leqq N_{2\nu-2}$ we conclude $N_{2\nu-3}\leqq 2\nu$. Further, for small values of t

$$R_{2\nu-3}(t) = -\frac{t^2}{2} + A_0 t > 0,$$

and

$$R_{2\nu-3}\left(\frac{1}{2}\right) = -\frac{1}{8} + \frac{1}{2}\sum_{i=0}^{(n-1)/2} A_i - \frac{1}{n}\sum_{i=0}^{(n-1)/2} iA_i = \frac{1}{8} - \frac{\sum_{i=0}^{(n-1)/2} iA_i}{n} < 0$$

because

$$\sum_{i=0}^{(n-1)/2} iA_i > \frac{n}{8}$$

in case $n\equiv 3 \pmod 4$ and $n\geqq 11$. Thus $\frac{1}{2}N_{2\nu-3}$ is an odd number, hence $N_{2\nu-3}\leqq 2\nu-2$. From the inequality

$$N_0 + 2\nu - 3 \leqq N_{2\nu-3} \leqq 2\nu - 2$$

we conclude again $N_0=1$, $N_{2\nu-3}=2\nu-2$. This proof does not apply to $n=2$, 4, 6, 3, 5, 7, 9. But the truth of the Lemma in these particular cases can be verified directly. This verification is not very laborious, since everything reduces to testing whether certain algebraic equations have roots between 0 and 1, and in all cases this is easily decided by application of the well known and simple Laguerre criterion.

9. Now that the equality $\alpha_\nu=1$ is proved it follows immediately that $\beta_s=0$ for $s\geqq\nu$, that is, *$G_{2s}(t)$ does not change sign in the interval* $0<t<1$. It suffices now to make use of formula (15) and apply repeatedly the integration by parts in (12) to obtain the following expansion of R_n:

$$R_n = \sum_{s=0}^{k-1} c_s\left\{f^{(2\nu+2s-1)}(1) - f^{(2\nu+2s-1)}(0)\right\} + c_k f^{(2\nu+2k)}(\xi), \tag{16}$$

where

$$c_s = -\frac{1}{(2\nu+2s)!}\sum_{i=0}^{n} A_i B_{2\nu+2s}\left(\frac{i}{n}\right) = -\frac{\gamma_{2\nu+2s}}{(2\nu+2s)!} \tag{17}$$

and ξ is an unknown number between 0 and 1. To show that this expansion possesses all the properties of the Euler-Maclaurin expansion it remains to prove that the numbers

$$\gamma_{2\nu},\ \gamma_{2\nu+2},\ \gamma_{2\nu+4},\ \cdots$$

alternate in sign. But this is almost evident if we notice that

$$\int_0^1 G_{2\nu+2s}(t)dt = -\gamma_{2\nu+2s}$$

and that $G_{2\nu+2s}(t)$ has the same sign as

$$G''_{2\nu+2s}(0) = (2\nu+2s)(2\nu+2s-1)\gamma_{2\nu+2s-2}.$$

It is easy to see that c_0 is negative, so that in general

$$(-1)^{s-1}c_s > 0.$$

The expansion (16) is especially useful when all the derivatives of an even order $\geqq 2\nu$ have the same sign in (0, 1). For then, retaining a certain number of terms in (16), the error in estimating R_n will be less in absolute value than the first neglected term and of the same sign.

10. Perhaps the simplest way of calculating coefficients c_s is to use their

expression (17). Since values of $B_{2s}(x)$ can be expressed simply through Bernoullian numbers for $x=\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{6}$, we have for $n=2, 3, 4, 6$ very elegant expansions:

$$\int_0^1 f(x)dx = \frac{1}{6}\left\{f(0) + 4f\left(\frac{1}{2}\right) + f(1)\right\}$$

$$+ \sum_{k=2,3,\cdots} \frac{(-1)^{k-1}(1-2^{2-2k})}{3} \frac{B_k}{(2k)!}\{f^{(2k-1)}(1) - f^{(2k-1)}(0)\},$$

$$\int_0^1 f(x)dx = \frac{1}{8}\left\{f(0) + 3f\left(\frac{1}{3}\right) + 3f\left(\frac{2}{3}\right) + f(1)\right\}$$

$$+ \sum_{k=2,3,\cdots} \frac{(-1)^{k-1}(1-3^{2-2k})}{8} \frac{B_k}{(2k)!}\{f^{(2k-1)}(1) - f^{(2k-1)}(0)\},$$

$$\int_0^1 f(x)dx = \frac{1}{90}\left\{7f(0) + 32f\left(\frac{1}{4}\right) + 12f\left(\frac{1}{2}\right) + 32f\left(\frac{3}{4}\right) + 7f(1)\right\}$$

$$+ \sum_{k=3,4,\cdots} (-1)^k \frac{1 - 5\cdot 2^{2-2k} + 4\cdot 4^{2-2k}}{45} \frac{B_k}{(2k)!}\{f^{(2k-1)}(1) - f^{(2k-1)}(0)\},$$

$$\int_0^1 f(x)dx = \frac{1}{840}\left\{41f(0) + 216f\left(\frac{1}{6}\right) + 27f\left(\frac{1}{3}\right) + 272f\left(\frac{1}{2}\right)\right.$$

$$\left. + 27f\left(\frac{2}{3}\right) + 216f\left(\frac{5}{6}\right) + 41f(1)\right\}$$

$$+ \sum_{k=4,5,\cdots} (-1)^{k-1} \frac{1 - 7\cdot 2^{4-2k} + 7\cdot 3^{4-2k} - 6^{4-2k}}{840} \frac{B_k}{(2k)!}$$

$$\times \{f^{(2k-1)}(1) - f^{(2k-1)}(0)\}.$$

The following table gives values of $\gamma_{2\nu}, \gamma_{2\nu+2}, \gamma_{2\nu+4}, \gamma_{2\nu+6}$ for $n=5, 7, 8$:

n	$\gamma_{2\nu}$	$\gamma_{2\nu+2}$	$\gamma_{2\nu+4}$	$\gamma_{2\nu+6}$
5	$\frac{11}{2^2\cdot 3\cdot 7\cdot 5^4}$	$-\frac{7}{2\cdot 3\cdot 5^5}$	$\frac{15351}{2^2\cdot 11\cdot 5^8}$	$-\frac{64427}{3\cdot 7\cdot 13\cdot 5^7}$
7	$\frac{167}{2\cdot 3^2\cdot 5\cdot 7^6}$	$-\frac{2665}{2\cdot 3\cdot 11\cdot 7^7}$	$\frac{1387331}{3\cdot 5\cdot 13\cdot 7^9}$	$-\frac{103112581}{3^2\cdot 7^{12}}$
8	$\frac{37}{2^{19}\cdot 3\cdot 11}$	$-\frac{235873}{2^{24}\cdot 3\cdot 5\cdot 7\cdot 13}$	$\frac{134671787}{2^{31}\cdot 3^2\cdot 5}$	$-\frac{180237358327}{2^{32}\cdot 3^2\cdot 5\cdot 17}$

It might be of interest for practical computers to know that the remainder in the much used G. F. Hardy's formula can be expanded in the same manner. We have in fact

$$\int_0^1 f(x)dx \doteq \frac{1}{3}\left[0.14f(0)+0.81f\left(\frac{1}{6}\right)+1.1f\left(\frac{1}{2}\right)+0.81f\left(\frac{5}{6}\right)+0.14f(1)\right]$$

$$+\frac{1}{21772800}\{f^{(5)}(1)-f^{(5)}(0)\}$$

$$+\sum_{k=4,5,6,\cdots}(-1)^{k-1}\frac{1-29\cdot 2^{1-2k}+81\cdot 3^{1-2k}-81\cdot 6^{1-2k}}{300}$$

$$\times\frac{B_k}{(2k)!}\{f^{(2k-1)}(1)-f^{(2k-1)}(0)\}.$$

Stanford University, Calif.

ON THE ASYMPTOTIC SOLUTIONS OF ORDINARY DIFFERENTIAL EQUATIONS, WITH REFERENCE TO THE STOKES' PHENOMENON ABOUT A SINGULAR POINT*

BY
RUDOLPH E. LANGER

1. **Introduction.** If in an ordinary differential equation of the second order, written in the form

$$(1) \qquad \frac{d^2w}{ds^2} + \{\lambda\psi(s) + \tau(\lambda, s)\}w = 0,$$

λ represents a large parameter, it is frequently of importance to know the character of the asymptotic dependence of the solutions upon this parameter and upon the variable s. The literature of differential equations records many investigations of this matter. If the variable s ranges over a real interval R_s, or more generally over a region R_s of the complex plane, on which the coefficients $\psi(s)$ and $\tau(\lambda, s)$ are bounded and the former is bounded from zero, there exist continuous forms composed of elementary functions of which each represents a solution over the entire region R_s.† On the other hand, if the coefficient $\psi(s)$ becomes zero at some point of R_s the situation is more intricate. To represent one and the same solution an asymptotic form must then be constructed of other than elementary functions, or in the alternative, i.e., if it is to be of the simpler type, it is subject to the Stokes' phenomenon. The latter requires that the form change abruptly in a specifiable but intricate way as certain frontiers both in the s and λ planes are traversed. The theory of these asymptotic solutions, it being still supposed that the coefficient $\tau(\lambda, s)$ is bounded as to s, has been given,‡ and applies to a number of standard differential equations. The list includes among others the equations for the Bessel functions,§ the Hermite or Weber functions,‖ the Mathieu func-

* Presented to the Society, September 6, 1934; received by the editors July 6, 1934.

† With certain conditions when R_s is infinite.

‡ Langer, R. E., *On the asymptotic solutions of ordinary differential equations*, etc., these Transactions, vol. 33 (1931), pp. 23–64, vol. 34 (1932), pp. 447–480, vol. 36 (1934), pp. 90–106. For a descriptive account, literature and applications cf. also Langer, R. E., *The asymptotic solutions of ordinary linear differential equations of the second order, with special reference to the Stokes' phenomenon*, Bulletin of the American Mathematical Society, vol. 40 (1934), pp. 545–582.

§ Langer, R. E., loc. cit.

‖ Schwid, N., *On the Asymptotic Forms of the Hermite and Weber Functions*, Thesis (1934) University of Wisconsin; see these Transactions, vol. 37 (1935), pp. 339–362.

tions,* the Laguerre functions and certain of the confluent forms of the hypergeometric functions.

The incidence of the Stokes' phenomenon has been associated with the vanishing of the coefficient $\psi(s)$, and its quantitative aspects with the order to which that coefficient becomes zero. It is shown in the present paper that this phenomenon is engendered also by an infinity in either of the coefficients, and is quantitatively dependent upon the structure of that infinity. More generally $\tau(\lambda, s)$ and $\psi(s)$ may, simultaneously or singly, become respectively infinite and either infinite or zero, and it is this inclusive situation which is discussed in the present investigation. Specifically the coefficient $\tau(\lambda, s)$ is admitted to have a pole of the first or second order, while $\psi(s)$ is taken to contain as a factor the quantity $(s-s_0)^\nu$, with ν a (any) real constant exceeding -2. A number of standard differential equations may be brought under this general type, in particular the equations for the ordinary or the associated Legendre functions, for the Laguerre functions, and for the Mathieu functions of higher order.† The asymptotic representations of such functions with variable and parameter complex, may accordingly be obtained by suitable specializations of the formulas with which the present discussion culminates.

2. **The normal form of the given differential equation.** Let the differential equation, reduced by the usual removal of the term of the first order, be of the form (1), with λ a sufficiently large value which may be complex. This equation is to be considered in a domain R_s which is simply connected; which may be finite or infinite; and which contains the point designated below by s_0. The equation, the domain, and the admitted range of parameter values are aggregately to fulfill a set of hypotheses which will be numbered from (i) to (vi) and which will be enunciated at appropriate points in the sequel. The initial pair are as follows:

(i) *Within R_s the coefficient $\psi(s)$ is of the form*

$$\psi(s) \equiv (s - s_0)^\nu \psi_1(s),$$

with $\nu > -2$, and with $\psi_1(s)$ a non-vanishing single-valued analytic function.

(ii) *Within R_s the coefficient $\tau(\lambda, s)$ is of the form*

$$\tau(\lambda, s) \equiv \frac{A_1}{(s - s_0)^2} + \frac{B_1}{s - s_0} + C_1(\lambda, s),$$

with A_1 and B_1 any constants, and $C_1(\lambda, s)$ an analytic function which on any finite portion of R_s is bounded uniformly with respect to λ.

* Langer, R. E., *The solutions of the Mathieu equation*, etc., these Transactions, vol. 36 (1934), pp. 637–695.

† Cf. Humbert, P., Proceedings of the Edinburgh Mathematical Society, 1921–22, p. 27.

In the product $\lambda\psi(s)$ the constant factors may be, and will be, taken to be distributed between λ and $\psi(s)$ so that the latter expands in the form

$$\psi(s) \equiv (s - s_0)^\nu \{1 + \alpha_1(s - s_0) + \alpha_2(s - s_0)^2 + \cdots \}.$$

If in this formula and in that for $\tau(\lambda, s)$ above, the constants α_1 and B_1 are both zero the differential equation will be defined to be *normal*. In this case at most a change of origin and lettering may be made to give it the form

$$\frac{d^2u}{dz^2} + \left\{\rho^2\phi^2(z) + \frac{\frac{1}{4} - A^2}{z^2} + \chi(\rho, z)\right\}u = 0. \tag{2}$$

In the contrary case the equation may always be normalized by the substitutions

$$(s - s_0) = \frac{z^2}{4}, \qquad w = z^{1/2}u,$$

the symbols of the form (1) leading then to those of (2) as is shown thus:

$$\rho = \lambda^{1/2}/2^{\nu+1}, \qquad A^2 = 1 - 4A_1,$$

$$\chi(\rho, z) = B_1 - \frac{z^2}{4} C_1(\lambda, s),$$

$$\phi^2(z) = z^{1/\mu - 2}\psi_1(s), \quad \mu = \frac{1}{2(\nu + 2)}.$$

The form (2) will be made basic for the discussion which follows. The facts to be especially noted at this point are the following. First, the constant μ is always real and positive (not zero), but is not otherwise restricted as to magnitude. Second, A^2 is an unrestricted constant real or complex. For definiteness the designation A will be reserved for that root of A^2 for which (unless it is zero)

$$-\pi/2 < \arg A \leqq \pi/2.$$

Third, if the constant ν is not an even integer, the region R_s must be considered as a Riemann surface with a branch point at s_0 if unique values are to be assigned to $\psi^{1/2}(s)$. The relation between s and z maps this region upon a domain, to be denoted by R_z, which is in general also a Riemann surface (or a part of such). Upon this surface, whose branch point is at the origin, the functions $\phi(z)$ and $\chi(\rho, z)$ are single-valued and analytic, and on any finite portion of it $\chi(\rho, z)$ is bounded uniformly with respect to the parameter ρ. Finally, the symbol $\phi(z)$ will be understood to represent that root of $\phi^2(z)$ which is determined by the relation

$$\lim_{z\to 0} \left\{\phi(z)/z^{1/(2\mu)-1}\right\} = 1.$$

3. **The "related" differential equation.** The formulas

$$\Phi = \int_0^z \phi(z)dz, \qquad \xi = \rho\Phi \tag{3}$$

may be looked upon as defining the complex variables Φ and ξ, whose domains of variation are to be designated respectively by R_Φ and R_ξ. Each of these domains in general lies upon a Riemann surface with branch point at the origin, and consists of sheets in number finite or infinite as the character of the constant μ may determine. In a neighborhood of the origin the correspondence between points Φ and z is of the "one to one" type. It is to be an hypothesis that this is so for the entire domains considered, i.e.,

(iii) *The region R_z is such that the correspondence of points of the Riemann surfaces R_Φ and R_z is unique.*

It will be observed that the domains R_ξ and R_Φ differ only in scale and orientation, the respective factors depending upon ρ. In particular, it is to be noted that since some neighborhood of $z=0$ lies in R_z, therefore the region $|\xi| \leqq N$, with any fixed constant N, lies entirely within R_ξ when $|\rho|$ is sufficiently large. Also as $\arg \rho$ varies the domain R_ξ rotates, so that any locus in ξ, fixed relatively to R_ξ, has an image locus in R_z which varies with ρ.

It is a consequence of the hypotheses that both ϕ and Φ are finite and different from zero except possibly at the origin. Hence the function

$$\Psi(z) \equiv \phi^{-1/2}(z)\Phi^{1/2-\mu}(z), \tag{4}$$

which is indeterminate at the origin, may be taken as so defined there that it and its reciprocal are analytic over the entire region R_z.

Consider the function

$$y(z) \equiv \Psi(z)\xi^{\mu}C_\beta(\xi), \tag{5}$$

in which C_β represents any cylinder function of the order β, the latter being as yet unspecified. It is found that this function solves the differential equation

$$y'' + \left\{\rho^2\phi^2 + (\mu^2 - \beta^2)\frac{\phi^2}{\Phi^2} - \frac{\Psi''}{\Psi}\right\}y = 0.$$

Since ϕ^2/Φ^2 differs from $1/(4\mu^2z^2)$ only by an analytic function, the choice

$$\beta = 2\mu A \tag{6}$$

reduces the equation to the form

$$\frac{d^2y}{dz^2} + \left\{\rho^2\phi^2(z) + \frac{\frac{1}{4} - A^2}{z^2} + \omega(z)\right\}y = 0, \tag{7}$$

with $\omega(z)$ a coefficient which is analytic throughout R_z. The differential equa-

tion (7), in so far as its more immediately essential features are concerned, is identical with the given equation (2). It will therefore be referred to as the *related* equation. It is explicitly and generally solved by the formula (5). The index β may evidently be zero or any complex constant subject to

$$-\pi/2 < \arg \beta \leqq \pi/2.$$

4. **The solutions of the related equation.** If the cylinder function in formula (5) is chosen in turn as the Bessel function of the first and second kind, the solutions obtained are

$$\begin{aligned} y_1(z) &\equiv \Psi(z)\xi^\mu J_\beta(\xi), \\ y_2(z) &\equiv \Psi(z)\xi^\mu Y_\beta(\xi). \end{aligned} \tag{8}$$

These functions are linearly independent, their Wronskian having the value

$$W(y_1, y_2) = \frac{-2}{\pi}\rho^{2\mu},$$

and near $\xi=0$ they are of the forms

$$\begin{aligned} y_1(z) &= \xi^{\mu+\beta}O(1), \\ y_2(z) &= \begin{cases} \xi^{\mu-\beta}O(1), & \text{if } \beta \neq 0, \\ \xi^\mu \log \xi O(1), & \text{if } \beta = 0, \end{cases} \end{aligned} \tag{9}$$

with $O(1)$ signifying in each case a bounded function.

From the formulas (9) it may be seen, since $\xi=O(z^{1/(2\mu)})$, that by an arbitrary approach to the origin on which arg z remains bounded the solution $y_1(z)$ invariably approaches zero. Moreover, if $\Re(A)>0$, $y_1(z)$ is determined uniquely (except for a constant factor) as the solution which vanishes at $z=0$ to a higher order than any other. By a similar approach to the origin $y_2(z)$ either also approaches zero or else becomes infinite according as the real component of the constant A is less than or greater than $\frac{1}{2}$. Clearly the same necessarily follows for any solution which is linearly independent of y_1.

In virtue of the relation

$$\frac{d\xi}{dz} = \frac{\rho^{2\mu}}{\Psi^2(z)}\xi^{1-2\mu}, \tag{10}$$

which is readily derived from (3) and (4), the differentiation of the formulas (9) with respect to z leads to

$$\begin{aligned} y_1'(z) &= \rho^{2\mu}\xi^{-\mu+\beta}O(1), \\ y_2'(z) &= \begin{cases} \rho^{2\mu}\xi^{-\mu-\beta}O(1), & \text{if } \beta \neq 0, \\ \rho^{2\mu}\xi^{-\mu} \log \xi O(1), & \text{if } \beta = 0. \end{cases} \end{aligned} \tag{9'}$$

The solutions (8) are especially adapted to such considerations as involve primarily values of ξ near the origin. On the other hand, when large values of ξ are in question the solutions obtained by using the Bessel functions of the third kind in formula (5) are of advantage. The formulas

$$
\begin{aligned}
y_{2m+1,1}(z) &\equiv y_{2m,1}(z), \\
y_{2m,1}(z) &\equiv \left(\frac{\pi}{2}\right)^{1/2} e^{(\beta+2m+1/2)\pi i/2}\Psi(z)\xi^{\mu}H_{\beta}^{(1)}(\xi e^{-2m\pi i}), \\
y_{2m-1,2}(z) &\equiv y_{2m,2}(z), \\
y_{2m,2}(z) &\equiv \left(\frac{\pi}{2}\right)^{1/2} e^{-(\beta+2m+1/2)\pi i/2}\Psi(z)\xi^{\mu}H_{\beta}^{(2)}(\xi e^{-2m\pi i})
\end{aligned} \tag{11}
$$

associate a pair of such solutions $y_{k,j}(z)$, $j=1, 2$, with each integral index k. The members of any pair are linearly independent, their Wronskian being

$$W(y_{k,1}, y_{k,2}) = 2i\rho^{2\mu},$$

and each is invariably independent of $y_1(z)$. Near the origin, therefore (if $\arg z$ is bounded),

$$
y_{k,j}(z) = \begin{cases} \xi^{\mu-\beta}O(1), & \text{if } \beta \neq 0, \\ \xi^{\mu}\log\xi O(1), & \text{if } \beta = 0, \end{cases} \qquad j = 1, 2, \tag{12}
$$

and

$$
y'_{k,j}(z) = \begin{cases} \rho^{2\mu}\xi^{-\mu-\beta}O(1), & \text{if } \beta \neq 0, \\ \rho^{2\mu}\xi^{-\mu}\log\xi O(1), & \text{if } \beta = 0. \end{cases} \tag{12'}
$$

Let ϵ be chosen as any fixed sufficiently small but positive constant. Then the domain R_ξ is sub-divided into a set of overlapping sub-regions $\Xi^{(l)}$, $l=0, \pm 1, \pm 2, \cdots$, by the relations

$$\Xi^{(l)}: \quad (l - 1 + \epsilon)\pi \leqq \arg \xi \leqq (l + 1 - \epsilon)\pi. \tag{13}$$

The corresponding sub-regions of R_z may without confusion be designated by the same symbol. Since the sub-region $\Xi^{(l)}$ on R_ξ is fixed relative to R_ξ, the sub-region $\Xi^{(l)}$ in R_z will be dependent upon and variable with ρ.

When ξ is numerically sufficiently large, a condition which is to be indicated briefly by the symbolism $|\xi|>N$, the solutions of the differential equation (7) admit of asymptotic representations. The solutions (11) were especially chosen so that their representations are peculiarly simple when ξ lies in suitably associated sub-regions (13), i.e., specifically

$$
(14)\qquad
\begin{aligned}
y_{2m,1}(z) &= \Psi(z)\xi^{\mu-1/2}e^{i\xi}\left\{1 + O\left(\frac{1}{\xi}\right)\right\}, \quad \text{for } \xi \text{ in } \Xi^{(2m)} \text{ and } \Xi^{(2m+1)},\\
y_{2m,2}(z) &= \Psi(z)\xi^{\mu-1/2}e^{-i\xi}\left\{1 + O\left(\frac{1}{\xi}\right)\right\}, \quad \text{for } \xi \text{ in } \Xi^{(2m-1)} \text{ and } \Xi^{(2m)}.
\end{aligned}
$$

The differentiation of these forms is permissible and yields

$$
(14')\qquad
\begin{aligned}
y'_{2m,1}(z) &= \frac{i\rho^{2\mu}}{\Psi(z)}\,\xi^{-\mu+1/2}e^{i\xi}\left\{1 + O\left(\frac{1}{\xi}\right)\right\}, \quad \text{for } \xi \text{ in } \Xi^{(2m)} \text{ and } \Xi^{(2m+1)},\\
y'_{2m,2}(z) &= \frac{-\,i\rho^{2\mu}}{\Psi(z)}\,\xi^{-\mu+1/2}e^{-i\xi}\left\{1 + O\left(\frac{1}{\xi}\right)\right\}, \quad \text{for } \xi \text{ in } \Xi^{(2m-1)} \text{ and } \Xi^{(2m)}.
\end{aligned}
$$

A solution of the differential equation is generally found to become exponentially infinite with $|\mathfrak{I}(\xi)|$. When such is the case the solution will be described as of the *dominant* type for the range of values ξ concerned. Exceptionally, however, the solution approaches zero under the stated circumstances. In that case it will be described as of the *sub-dominant* type. From the formulas (14) it is evident that in the domain common to the sub-regions $\Xi^{(2m)}$ and $\Xi^{(2m+1)}$ the associated solution $y_{2m,1}(z)$ is sub-dominant, since in this domain $\mathfrak{I}(\xi)\to+\infty$ with $|\xi|$. In the remaining parts of the specified sub-regions it is dominant. Likewise the solution $y_{2m,2}(z)$ is sub-dominant in the domain common to $\Xi^{(2m-1)}$ and $\Xi^{(2m)}$, since there $\mathfrak{I}(\xi)\to-\infty$, and in the remaining parts of these sub-regions is dominant. These facts may be briefly though loosely stated thus: *For any index k, $y_{k,1}(z)$ is sub-dominant in the "upper half" and dominant in the "lower half" of $\Xi^{(k)}$, while $y_{k,2}(z)$ is sub-dominant in the "lower half" and dominant in the "upper half" of $\Xi^{(k)}$.* Since any solution which is linearly independent of the sub-dominant one on a given range must include a component of the dominant one, it must evidently itself be of the dominant type. It follows that the solutions (11) are in fact determined uniquely (except for constant factors) by their properties of sub-dominance as described.

The formulas (14), specifying as they do certain sub-regions (13), in general cease to be valid when ξ passes out of these regions. This is the Stokes' phenomenon. To obtain the representations when ξ lies in some other non-associated sub-region, say $\Xi^{(h)}$, it is, however, merely necessary to express the given solutions in terms of the solutions $y_{h,j}(z)$, and to utilize the formulas (14) for the latter. The relation between distinct pairs of solutions (11) which is thus brought into question is obtainable from the known relation

between the Bessel functions involved.* It is found in this way that when $|\xi|>N$, then

$$
(15)\qquad \begin{aligned} y_{2m,j}(z) = \Psi(z)\xi^{\mu-1/2}\Big\{ & \overset{(m)}{c_{j,1}} e^{i\xi}\Big[1 + O\Big(\frac{1}{\xi}\Big)\Big] \\ & + \overset{(m)}{c_{j,2}} e^{-i\xi}\Big[1 + O\Big(\frac{1}{\xi}\Big)\Big]\Big\}, \end{aligned}
$$

with coefficients as follows:

$$
(16)\qquad \begin{aligned} &\left.\begin{aligned} \overset{(m)}{c_{1,1}} &= (-1)^{m-s+1}\frac{\sin(2s-2m-1)\beta\pi}{\sin\beta\pi} \\ \overset{(m)}{c_{2,1}} &= (-1)^{m-s+1}\frac{i\sin(2s-2m)\beta\pi}{\sin\beta\pi} \end{aligned}\right\} \text{ for } \xi \text{ in } \Xi^{(2s-1)} \text{ or } \Xi^{(2s)}, \\ &\left.\begin{aligned} \overset{(m)}{c_{1,2}} &= (-1)^{m-s+1}\frac{i\sin(2s-2m)\beta\pi}{\sin\beta\pi} \\ \overset{(m)}{c_{2,2}} &= (-1)^{m-s}\frac{\sin(2s-2m+1)\beta\pi}{\sin\beta\pi} \end{aligned}\right\} \text{ for } \xi \text{ in } \Xi^{(2s)} \text{ or } \Xi^{(2s+1)}. \end{aligned}
$$

The asymptotic form of the solution $y_1(z)$ is analogously given by the formulas†

$$
(17)\qquad y_1(z) = \Psi(z)\xi^{\mu-1/2}\Big\{c_1e^{i\xi}\Big[1 + O\Big(\frac{1}{\xi}\Big)\Big] + c_2e^{-i\xi}\Big[1 + O\Big(\frac{1}{\xi}\Big)\Big]\Big\},
$$

and

$$
(17')\qquad y_1'(z) = \frac{i\rho^{2\mu}\xi^{-\mu+1/2}}{\Psi(z)}\Big\{c_1e^{i\xi}\Big[1 + O\Big(\frac{1}{\xi}\Big)\Big] - c_2e^{-i\xi}\Big[1 + O\Big(\frac{1}{\xi}\Big)\Big]\Big\},
$$

with coefficients

$$
(18)\qquad \begin{aligned} c_1 &= (2\pi)^{-1/2}e^{(2s-1/2)(\beta+1/2)\pi i}, \quad \text{for } \xi \text{ in } \Xi^{(2s-1)} \text{ or } \Xi^{(2s)}, \\ c_2 &= (2\pi)^{-1/2}e^{(2s+1/2)(\beta+1/2)\pi i}, \quad \text{for } \xi \text{ in } \Xi^{(2s)} \text{ or } \Xi^{(2s+1)}. \end{aligned}
$$

Since the coefficients (18) are different from zero for every index, the solution $y_1(z)$ is seen to be of the dominant type in both the upper and the lower "half" of any and every sub-region (13).

For subsequent use it may be observed that the expression

$$
\{y_p(z)y_q(z_1) - y_q(z)y_p(z_1)\}/W(y_p, y_q),
$$

looked upon as a ratio of determinants, is obviously independent of the choice of the solutions y_p and y_q. It follows that the expression

* Watson, G. M., *A Treatise on the Theory of Bessel Functions*, Cambridge, 1922, p. 75.

† Watson, loc. cit., p. 202.

$$(19a)\quad Q(z, z_1) \equiv \frac{\pi}{2}\{y_1(z)y_2(z_1) - y_2(z)y_1(z_1)\}\frac{\Psi^3(z_1)}{\Psi(z)}\{\chi(\rho, z_1) - \omega(z_1)\}\xi_1^{2\mu-1},$$

which is to be used below, may equally well be written

$$(19b)\quad Q(z, z_1) \equiv \frac{i}{2}\{y_{k,1}(z)y_{k,2}(z_1) - y_{k,2}(z)y_{k,1}(z_1)\}\frac{\Psi^3(z_1)}{\Psi(z)}\{\chi(\rho, z_1) - \omega(z_1)\}\xi_1^{2\mu-1},$$

with any choice of the index k.

5. **The solution $u_1(z)$ when $|\xi| \leqq N$.** It is convenient for the considerations at hand to designate briefly as a "Γ curve" any *ordinary curve upon which as seen in R_ξ the ordinate varies monotonically with the arc length, and upon which the variation of* arg ξ *remains below a (some) finite bound independent of the particular curve.* Inasmuch as this description is relative to R_ξ the curves in question as seen in R_z depend, of course, upon ρ. The following, which is to be made an hypothesis, therefore essentially applies to a conjunction of the admitted range of values ρ with the configuration of the region R_z.

(iv) *The region R_z is such that for any (every) admitted value of ρ each point may be connected with the origin by some "Γ curve" which lies entirely in the region.*

Let the function $\theta(\rho, z)$ be defined by the formula

$$\theta(\rho, z) \equiv \chi(\rho, z) - \omega(z).$$

It is clearly analytic in z and bounded uniformly as to ρ in any finite portion of R_z. Since the differential equation (2) may be written in the form

$$\frac{d^2u}{dz^2} + \left\{\rho^2\phi^2(z) + \frac{\frac{1}{4} - A^2}{z^2} + \omega(z)\right\}u = -\theta(\rho, z)u,$$

the left member of which is identical in structure with that of the equation (7), it follows that the formula

$$(20)\quad u(z) = y(z) + \frac{\pi}{2\rho^{2\mu}}\int_{z_0}^{z}\{y_1(z)y_2(z_1) - y_2(z)y_1(z_1)\}\theta(\rho, z_1)u(z_1)dz_1,$$

with any limit z_0 independent of z, relates a solution of equation (2) with any solution $y(z)$ of the equation (7). This relationship will be indicated consistently by the use of similar subscripts.

The differentiation of (20) yields the associated formula

$$(20')\quad u'(z) = y'(z) + \frac{\pi}{2\rho^{2\mu}}\int_{z_0}^{z}\{y_1'(z)y_2(z_1) - y_2'(z)y_1(z_1)\}\theta(\rho, z_1)u(z_1)dz_1.$$

The path of integration, to be inferred from (20) or (20′) as in R_z, may with greater convenience be considered in the region R_ξ, the transformation being facilitated by the use of formula (10). If the fixed limit of integration is chosen as the origin, and this is to be the case in the present and the following section, the path may and will be chosen as a curve of the type Γ. With the introduction of the abbreviations

$$(21) \qquad Y_1(z) \equiv \frac{e^{-i\xi}}{\Psi(z)}\, y_1(z), \quad U_1(z) \equiv \frac{e^{-i\xi}}{\Psi(z)}\, u_1(z),$$

the relation (20) may then be written in the form

$$U_1(z) = Y_1(z) + \frac{1}{\rho^{4\mu}} \int_\Gamma Q(z, z_1) e^{-i(\xi-\xi_1)} U_1(z_1) d\xi_1,$$

with $Q(z, z_1)$ the quantity defined in (19a). This is an integral equation for $U_1(z)$. By the familiar process of successive iteration it leads formally to the relation

$$(22) \qquad U_1(z) = Y_1(z) + \sum_{n=1}^{\infty} Y_1^{(n)}(z),$$

with

$$(23) \qquad Y_1^{(n+1)}(z) = \frac{1}{\rho^{4\mu}} \int_\Gamma Q(z, z_1) e^{-i(\xi-\xi_1)} Y_1^{(n)}(z_1) d\xi_1,$$

and $Y_1^{(0)}(z) \equiv Y_1(z)$. Whenever it is uniformly convergent the relation (22) is a true formula for $U_1(z)$. It is to be shown that this is the case whenever $|\rho|$ is sufficiently large.

Consider the relation

$$(24) \qquad Y_1^{(n)}(z) = \frac{\xi^{4n\mu+\mu+\beta}}{\rho^{4n\mu}} O(1), \quad \text{when } |\xi| \leqq N,$$

with $O(1)$ representing a function which is bounded uniformly as to n. The relation is evidently satisfied when $n=0$ because of the formulas (9) and the boundedness of $|\xi|$. It may be shown as follows, however, that the validity of the relation for any n implies it for the next larger value, so that by induction the relation will be established for all n.

Since the region $|\xi| \leqq N$ lies entirely in R_ξ the path of integration in (23) may be taken straight. Then from the formulas (9) and (19a) it is seen that if $\beta \neq 0$ the character of the relation (23) is

$$Y_1^{(n+1)} = \frac{1}{\rho^{4\mu+4n\mu}} \int_\Gamma \left\{ \xi^{\mu+\beta} \xi_1^{4(n+1)\mu} O(1) + \xi^{\mu-\beta} \xi_1^{4(n+1)\mu+2\beta} O(1) \right\} \frac{d\xi_1}{\xi_1}.$$

But on setting $t=\xi_1/\xi$ this may be written

$$Y_1^{(n+1)} = \frac{\xi^{4(n+1)\mu+\mu+\beta}}{\rho^{4(n+1)\mu}} \int_0^1 \{O(1) + t^{2\beta}O(1)\} t^{4(n+1)\mu-1}dt,$$

and since μ is positive and the real part of β is not negative this establishes (24) with n replaced by $n+1$. If $\beta=0$ it is found in a similar way that

$$Y_1^{(n+1)} = \frac{1}{\rho^{4\mu+4n\mu}} \int_\Gamma \xi^\mu \xi_1^{4(n+1)\mu} \{\log \xi_1 - \log \xi\} O(1) \frac{d\xi_1}{\xi_1},$$

and since this may be written

$$Y_1^{(n+1)} = \frac{\xi^{4(n+1)\mu+\mu}}{\rho^{4(n+1)\mu}} \int_0^1 t^{4(n+1)\mu-1} \log t\, O(1)dt,$$

the conclusion again follows. The relation (24) is therefore generally valid.

It is evident now, in virtue of (24), that the uniform convergence of the series in (22) over the region $|\xi| \leqq N$ is assured when $|\rho|$ is sufficiently large. Since that is a blanket assumption for the entire discussion it follows that

$$U_1(z) = Y_1(z) + \frac{\xi^{5\mu+\beta}}{\rho^{4\mu}} O(1), \quad \text{when } |\xi| \leqq N,$$

i.e., more explicitly

$$u_1(z) = \Psi(z)\xi^{\mu+\beta}\left\{\xi^{-\beta}J_\beta(\xi) + \frac{\xi^{4\mu}O(1)}{\rho^{4\mu}}\right\} \quad \text{when } |\xi| \leqq N. \tag{25}$$

The formula (25) describes the solution $u_1(z)$ near the origin. Since it is of the order of $\xi^{\mu+\beta}$, i.e., of the order of $z^{1/2+A}$, while the exponents of the differential equation relative to $z=0$ are $\frac{1}{2}\pm A$, it is seen that when $\Re(A)>0$, $u_1(z)$ is the solution which vanishes at the origin to a higher order than any which is linearly independent of it. It may also be observed in connection with the formula (25) that the second term, i.e., the vague correction term, is of the order of z^2 as well as of the order of $\rho^{-4\mu}$ relative to the explicit first term.

Finally, the substitution of the values (25) and (9′) in the right-hand member of the derived relation (20′) yields for the integral concerned a form

$$\frac{\xi^{3\mu+\beta}}{\rho^{2\mu}} \int_0^1 t^{4\mu-1}\{O(1) + t^{2\beta}O(1)\}dt.$$

Hence it may be concluded in precisely the manner above that

$$u_1'(z) = y_1'(z) + \frac{\xi^{3\mu+\beta}O(1)}{\rho^{2\mu}}, \quad \text{when } |\xi| \leqq N. \tag{25'}$$

This formula is precisely that which is obtained by a direct formal differentiation of (25).

6. **The solution $u_1(z)$ when $|\xi| > N$.** When $|\xi|$ is large the first or the second term of the formula (17) for $y_1(z)$ is dominant according as $\mathfrak{J}(\xi)$ is negative or positive. The deductions for the solution $u_1(z)$ are to be based largely upon this formula and must, therefore, be appropriately adapted to the location of ξ. Since such adaptation extends merely to formal and rather obvious detail the explicit argument will be given only say for $\mathfrak{J}(\xi) \leqq 0$. The function $Y_1(z)$ defined in (21) is then of the structure $\xi^{\mu-1/2}O(1)$ when $|\xi| > N$,* the same being true moreover of the functions defined by the formulas

$$Y_{k,1}(z) \equiv \frac{e^{-i\xi}}{\Psi(z)}\, y_{k,1}(z), \qquad Y_{k,2}(z) \equiv \frac{e^{i\xi}}{\Psi(z)}\, y_{k,2}(z), \tag{26}$$

provided k is the index of the sub-region (13) in which ξ is located, i.e.,

$$Y_{k,j}(z) = \xi^{\mu-1/2}O(1), \quad j = 1, 2, \text{when } |\xi| > N, \text{ and } \xi \text{ is in } \Xi^{(k)}. \tag{27}$$

Let ρ_μ be defined appropriately to the value of μ in the manner

$$\rho_\mu = \begin{cases} \rho, & \text{if } \mu > \frac{1}{4}, \\ \rho/\log \rho, & \text{if } \mu = \frac{1}{4}, \\ \rho^{4\mu}, & \text{if } \mu < \frac{1}{4}. \end{cases} \tag{28}$$

It evidently becomes infinite with ρ in every case. When $n = 0$ the relation

$$Y_1^{(n)}(z) = \frac{\xi^{\mu-1/2}}{\rho_\mu^n}\, O(1), \quad \text{when } |\xi| > N, \tag{29}$$

is valid, as was observed above. It is to be shown on the basis of the relation (23) that it is valid for all n.

For the consideration of the integral in (23) let the Γ curve of integration be sub-divided into the following component arcs: Γ_1, the arc on which $|\xi_1| \leqq N$; Γ_2, the arc on which $|\xi_1| > N$ but whose image in R_z lies within a (any specified) finite portion of R_z; Γ_3, the remaining arc if any. If R_z is finite no arc Γ_3 need be considered.

* The formula (21) was in fact designed to produce this result. If ξ is to be taken in an upper half-plane the formula (21) should be modified by replacing $-i$ by i, a change which in no way affects the reasoning in §5.

The relations (19b) and (10) may be used to show without difficulty that the following are equivalent formulas, i.e.,

$$(30)\quad Q(z, z_1)e^{-i(\xi-\xi_1)} = \begin{cases} \dfrac{i\Psi^3(z_1)\theta(\rho, z_1)}{2}\{Y_{k,1}(z)y_{k,2}(z_1) \\ \qquad - Y_{k,2}(z)y_{k,1}(z_1)e^{-2i\xi}\}e^{i\xi_1}\xi_1^{2\mu-1}, \\ \dfrac{i\Psi^4(z_1)\theta(\rho, z_1)}{2}\{Y_{k,1}(z)Y_{k,2}(z_1) \\ \qquad - Y_{k,2}(z)Y_{k,1}(z_1)e^{-2i(\xi-\xi_1)}\}\xi_1^{2\mu-1}, \\ \dfrac{i\theta(\rho, z_1)}{2\rho^{1-4\mu}\phi(z_1)}\{Y_{k,1}(z)Y_{k,2}(z_1) \\ \qquad - Y_{k,2}(z)Y_{k,1}(z_1)e^{-2i(\xi-\xi_1)}\}\xi_1^{1-2\mu}\dfrac{dz_1}{d\xi_1}. \end{cases}$$

They are to be used respectively for evaluating the integrations in (23) over the arcs Γ_1, Γ_2 and Γ_3. The exponential factors are then bounded, since $\Im(\xi-\xi_1)\leqq 0$ whenever ξ_1 is on a Γ curve which joins the origin with the point ξ. With the use of the formulas (27) when the variable is z; the formulas (12) and (24) when the variable is z_1 on Γ_1; and (27) and (29) when z_1 is on Γ_2 or Γ_3; it may then be shown that the relation (23) is structurally as follows:

$$Y_1^{(n+1)} = \frac{\xi^{\mu-1/2}}{\rho_\mu^{n+1}}\left\{\left(\frac{\rho_\mu}{\rho^{4\mu}}\right)^{n+1}\int_{\Gamma_1} O(1)d\xi_1 + \frac{\rho_\mu}{\rho^{4\mu}}\int_{\Gamma_2}\xi_1^{4\mu-1}O(1)\frac{d\xi_1}{\xi_1}\right.$$
$$\left. + \frac{\rho_\mu}{\rho}\int_{\Gamma_3}\frac{\theta(\rho, z_1)}{\phi(z_1)}O(1)dz_1\right\}.$$

In this the integral over Γ_1 and its coefficient are obviously bounded. The integral over Γ_2 is of the order of $\xi_1^{4\mu-1}$, $\log \xi_1$, or 1, according as 4μ is greater than, equal to, or less than 1. Since on this arc the value of ξ_1 is at most of the order of ρ, the order of the integral is seen to be the reciprocal of its coefficient so that the product of the two is bounded. The coefficient of the integral over Γ_3 is bounded. To insure finally the boundedness of this integral as well, the following is to be added as an hypothesis upon the given differential equation:

(v) *In the region R_z a relation*

$$\int\left|\frac{\theta(\rho, z)}{\phi(z)}dz\right| < M$$

is satisfied by some constant M, uniformly with respect to all arcs of integration which are of the type Γ for some admitted value of ρ, and on which $|z|\geqq N_1>0$.

The form of $Y_1^{(n+1)}$ has thus been shown to be as given by (29), and the latter, therefore, to be valid for all n. The convergence of the formula (22) when ρ is suitably large is therefore assured, whence it may be drawn that

$$U_1(z) = Y_1(z) + \frac{\xi^{\mu-1/2}}{\rho_\mu} O(1), \quad \text{when } |\xi| > N,$$

i.e., that

$$u_1(z) = y_1(z) + \Psi(z)\xi^{\mu-1/2}e^{i\xi}\frac{O(1)}{\rho_\mu}.$$

This is the result obtained on the assumption that ξ remains in a lower half-plane. The form correspondingly obtained when ξ is in an upper half-plane differs from it only in that the factor $e^{i\xi}$ is replaced by $e^{-i\xi}$. For unrestricted variation of ξ the result may, therefore, be expressed by the formula

$$(31)\quad u_1(z) = y_1(z) + \Psi(z)\xi^{\mu-1/2}\left\{\frac{e^{i\xi}O(1) + e^{-i\xi}O(1)}{\rho_\mu}\right\}, \quad \text{when } |\xi| > N.$$

A discussion entirely similar to that given but based upon the derived formula (20′) may be made to yield the representation of $u_1'(z)$, or alternatively the direct differentiability of (31) may be justified. Upon substituting for $y_1(z)$ its forms (17) and (17′), it is to be concluded that when $|\xi|>N$,

$$(32)\quad \begin{aligned} u_1(z) = \Psi(z)\xi^{\mu-1/2}\Big\{&c_1e^{i\xi}\left[1 + O\left(\frac{1}{\xi}\right) + O\left(\frac{1}{\rho_\mu}\right)\right] \\ &+ c_2e^{-i\xi}\left[1 + O\left(\frac{1}{\xi}\right) + O\left(\frac{1}{\rho_\mu}\right)\right]\Big\}, \end{aligned}$$

and

$$(32')\quad \begin{aligned} u_1'(z) = \frac{i\rho^{2\mu}\xi^{-\mu+1/2}}{\Psi(z)}\Big\{&c_1e^{i\xi}\left[1 + O\left(\frac{1}{\xi}\right) + O\left(\frac{1}{\rho_\mu}\right)\right] \\ &- c_2e^{-i\xi}\left[1 + O\left(\frac{1}{\xi}\right) + O\left(\frac{1}{\rho_\mu}\right)\right]\Big\}, \end{aligned}$$

with the coefficients (18).

THEOREM 1. *Under the hypotheses* (i) *to* (v) *the solution of the differential equation* (2) *with the exponent* $\frac{1}{2}+A$ *relative to the origin has the form* (25), (25′) *for values of* z *such that* $|\xi|\leq N$, *and the asymptotic forms* (32), (32′) *with coefficients* (18) *when* $|\xi|>N$.

7. The sub-dominant solutions. The results of the preceding section show that the solution $u_1(z)$ is of the dominant type for all admitted ranges of the variable. It is to be shown now that the differential equation admits

also solutions which are of the sub-dominant form in appropriately associated domains. These solutions are analogous to the solutions $y_{k,j}(z)$ of the related equation, and will, of course, be linearly independent of $u_1(z)$.

The domain common to a pair of adjacent sub-regions (13) lies wholly in either an upper or a lower half-plane. Its boundary, a portion of the boundary of R_ξ, either contains a point to be denoted by ξ_M, at which $\mathfrak{I}(\xi)$ is numerically a maximum, or else $\mathfrak{I}(\xi)$ is unbounded, the domain extending to infinity. A distinction between these cases will be avoided by permitting ξ_M to designate either the finite or the infinite point. There will be one such point for each pair of adjacent sub-regions (13). Inasmuch as these points are fixed relative to R_ξ the image points z_M in R_z in general vary with the parameter. The following hypothesis (the final one to be made) therefore again concerns the configurative character of R_z together with the admitted range of values ρ.

(vi) *The region R_z is such that for any (every) admitted value of ρ each point ξ_M may be connected with any (every) point of R_ξ in its respective half-plane by a curve of the type Γ lying entirely in R_ξ.*

It will be evident that in virtue of the earlier hypothesis (iv) each point ξ_M may be connected by a Γ curve not merely with any point of its own half-plane as stated, but in fact with any point of the two sub-regions (13) within which it lies.

A specific point ξ_M lies either above or below the axis of reals and some adaptation of the details to the case in point must be made. In principle, however, the argument is general and the explicit discussion will therefore be given only say for a case in which $\mathfrak{I}(\xi_M)<0$. The regions $\Xi^{(k)}$ in which ξ_M is included are then given by $k=2m-1$ and $k=2m$ with a suitable integer m. The variable ξ will be supposed to remain in these same sub-regions throughout the discussion of this section.

Let the formula (20) be written with the roles of $y(z)$, z_0, and $u(z)$ taken by $y_{k,2}(z)$, z_M, and $u_{k,2}(z)$, the path of integration being chosen as a Γ curve. With the use of the abbreviation

$$U_{k,2}(z) \equiv \frac{e^{i\xi}}{\Psi(z)}\, u_{k,2}(z),$$

the relation may then be written

$$\begin{aligned} U_{2m-1,2}(z) &\equiv U_{2m,2}(z), \\ U_{2m,2}(z) &\equiv Y_{2m,2}(z) + \frac{1}{\rho^{4\mu}} \int_\Gamma Q(z, z_1) e^{i(\xi-\xi_1)} U_{2m,2}(z_1) d\xi_1, \end{aligned}$$

and this by iteration leads to the formal relation

$$U_{2m,2}(z) = Y_{2m,2}(z) + \sum_{n=1}^{\infty} Y^{(n)}_{2m,2}(z), \tag{33}$$

with

$$Y^{(n+1)}_{2m,2}(z) = \frac{1}{\rho^{4\mu}} \int_{\Gamma} Q(z, z_1) e^{i(\xi-\xi_1)} Y^{(n)}_{2m,2}(z_1) d\xi_1, \tag{34}$$

and $Y^{(0)}_{m,2}(z) \equiv Y_{2m,2}(z)$. It is to be shown that for all n

$$Y^{(n)}_{2m,2}(z) = \frac{\xi^{\mu-1/2}}{\rho^n_\mu} O(1), \quad \text{when } |\xi| > N. \tag{35}$$

This is a fact when $n=0$, since the relation is then included in (27).

When ξ lies in the region $|\xi|>N$ the same may be assumed of the entire curve Γ,* and the latter therefore consists of at most arcs Γ_2 and Γ_3 as such were described in §6. Upon these arcs the second and third of the formulas (30) multiplied by $e^{2i(\xi-\xi_1)}$ may be respectively used to give the kernel of the formula (34). The exponentials involved are seen to be bounded since $\Im(\xi-\xi_1) \geqq 0$ when ξ_1 is on a Γ curve joining ξ with a point ξ_M such as is being considered. If then the relation (35) is assumed to hold for any n, the structure of (34) is

$$Y^{(n+1)}_{2m,2} = \frac{\xi^{\mu-1/2}}{\rho^{n+1}_\mu} \left\{ \frac{\rho_\mu}{\rho} \int_{\Gamma_3} \frac{\theta(\rho, z_1)}{\phi(z_1)} O(1) dz_1 + \frac{\rho_\mu}{\rho^{4\mu}} \int_{\Gamma_2} \xi_1^{4\mu-1} O(1) \frac{d\xi_1}{\xi_1} \right\},$$

and the members within the brace are bounded as was found in §6. The relation (35) is thus generally valid, the formula (33) accordingly convergent, and in consequence

$$U_{2m,2}(z) = Y_{2m,2}(z) + \frac{\xi^{\mu-1/2}}{\rho_\mu} O(1), \quad \text{when } |\xi| > N.$$

Hence

$$u_{2m-1,2}(z) \equiv u_{2m,2}(z),$$

$$u_{2m,2}(z) = y_{2m,2}(z) + \Psi(z)\xi^{\mu-1/2} e^{i\xi} \frac{O(1)}{\rho_\mu}, \quad \text{when } |\xi| > N, \tag{36a}$$

$$\text{and } \xi \text{ is in } \Xi^{(2m-1)} \text{ or } \Xi^{(2m)},$$

and the formula may be shown to be differentiable.

* The distortion of the curve which may be necessary to circumvent the domain $|\xi| \leqq N$ is always slight and is readily seen to be negligible.

When ξ_M lies in an upper half-plane it is contained in a pair of sub-regions $\Xi^{(2m)}$, $\Xi^{(2m+1)}$. The discussion corresponding to that given above and based on the abbreviation

$$U_{k,1}(z) \equiv \frac{e^{-i\xi}}{\Psi(z)} u_{k,1}(z)$$

is found then to lead to the result

$$\begin{aligned} &u_{2m+1,1}(z) \equiv u_{2m,1}(z), \\ \text{(36b)} \qquad &u_{2m,1}(z) = y_{2m,1}(z) + \Psi(z)\xi^{\mu-1/2}e^{i\xi}\frac{O(1)}{\rho_\mu}, \quad \text{when } |\xi| > N, \\ &\qquad\qquad \text{and } \xi \text{ is in } \Xi^{(2m)} \text{ or } \Xi^{(2m+1)}. \end{aligned}$$

On substituting the forms (14) it is thus established that when $|\xi|>N$

$$\text{(37)} \quad \begin{aligned} u_{2m,1}(z) &= \Psi(z)\xi^{\mu-1/2}e^{i\xi}\left\{1 + O\left(\frac{1}{\xi}\right) + O\left(\frac{1}{\rho_\mu}\right)\right\}, \\ u_{2m,1}'(z) &= \frac{i\rho^{2\mu}\xi^{-\mu+1/2}}{\Psi(z)} e^{i\xi}\left\{1 + O\left(\frac{1}{\xi}\right) + O\left(\frac{1}{\rho_\mu}\right)\right\}, \\ &\qquad\qquad \text{for } \xi \text{ in } \Xi^{(2m)} \text{ and } \Xi^{(2m+1)}; \\ u_{2m,2}(z) &= \Psi(z)\xi^{\mu-1/2}e^{-i\xi}\left\{1 + O\left(\frac{1}{\xi}\right) + O\left(\frac{1}{\rho_\mu}\right)\right\}, \\ u_{2m,2}'(z) &= \frac{-i\rho^{2\mu}\xi^{-\mu+1/2}}{\Psi(z)} e^{-i\xi}\left\{1 + O\left(\frac{1}{\xi}\right) + O\left(\frac{1}{\rho_\mu}\right)\right\}, \\ &\qquad\qquad \text{for } \xi \text{ in } \Xi^{(2m-1)} \text{ and } \Xi^{(2m)}. \end{aligned}$$

8. **The solutions $u_{k,j}(z)$ for unrestricted values of z.** When $|\xi| \leqq N$ the path of integration in the formula (34) may be chosen as a Γ curve from ξ_M to the edge of the domain $|\xi| \leqq N$, and thence to the point ξ as any ordinary curve on which $|\xi_1|$ decreases monotonically and $\arg \xi_1$ is bounded. The initial part will thus consist of at most arcs of type Γ_3 and Γ_2, and on these the equivalent forms

$$Q(z, z_1)e^{i(\xi-\xi_1)} = \begin{cases} \dfrac{i\theta(\rho, z_1)e^{i\xi}}{2\rho^{1-4\mu}\phi(z_1)}\{y_{k,1}(z)Y_{k,2}(z_1)e^{-2i\xi_1} - y_{k,2}(z)Y_{k,1}(z_1)\}\xi_1^{1-2\mu}\dfrac{dz_1}{d\xi_1}, \\ \dfrac{i\Psi^4(z_1)\theta(\rho, z_1)e^{i\xi}}{2\Psi(z)}\{y_{k,1}(z)Y_{k,2}(z_1)e^{-2i\xi_1} - y_{k,2}(z)Y_{k,1}(z_1)\}\xi_1^{2\mu-1} \end{cases}$$

may be respectively used in conjunction with the relations (35), (27) and (12). For the integration over the remaining arc, which will for convenience

be denoted by Γ_1, the formula (19a) may be drawn on for $Q(z, z_1)$, and used in conjunction with (9). If it is assumed then that the relation

$$(38)\qquad Y_{2m,2}^{(n)}(z) = \begin{cases} \dfrac{\xi^{\mu-\beta}}{\rho_\mu^n} O(1), & \text{if } \beta \neq 0, \\ \dfrac{\xi^\mu \log \xi}{\rho_\mu^n} O(1), & \text{if } \beta = 0, \end{cases} \quad \text{when } |\xi| \leqq N,$$

holds for any specific value of n, it is found that when $\beta \neq 0$ (34) is of the form

$$Y_{2m,2}^{(n+1)} = \frac{\xi^{\mu-\beta}}{\rho_\mu^{n+1}} \left\{ \frac{\rho_\mu}{\rho} \int_{\Gamma_3} \frac{\theta(\rho, z_1)}{\phi(z_1)} O(1) dz_1 + \frac{\rho_\mu}{\rho^{4\mu}} \int_{\Gamma_2} \xi_1^{4\mu-1} O(1) \frac{d\xi_1}{\xi_1} \right.$$
$$\left. + \frac{\rho_\mu}{\rho^{4\mu}} \int_{\Gamma_1} \left[\left(\frac{\xi}{\xi_1}\right)^{2\beta} O(1) + O(1) \right] \xi_1^{4\mu-1} d\xi_1 \right\}.$$

The formula when $\beta = 0$ differs from this in details which will now be familiar. The terms involving the integrals over Γ_3 and Γ_2 are of forms which have previously been discussed, and found to be bounded. In the integral over Γ_1 the quantity (ξ/ξ_1) never exceeds unity in numerical value and is of bounded argument, while $\Re(2\beta) \geqq 0$, and as a result the term in question is seen to be likewise bounded. Hence the relation (38) remains valid when n is replaced by $n+1$, and since it is evidently so when $n=0$ it is valid for all n.

From (38) and (33) it follows that the formula

$$U_{2m,j}(z) = Y_{2m,j}(z) + \begin{cases} \dfrac{\xi^{\mu-\beta}}{\rho_\mu} O(1), & \text{if } \beta \neq 0, \\ \dfrac{\xi^\mu \log \xi}{\rho_\mu} O(1), & \text{if } \beta = 0, \end{cases} \quad |\xi| \leqq N,$$

holds when $j=2$. An entirely similar discussion may be made to show that it holds also when $j=1$. The conclusion to be drawn is, therefore, that when $|\xi| \leqq N$ and arg z is bounded,

$$(39)\qquad u_{2m,j}(z) = y_{2m,j}(z) + \begin{cases} \Psi(z)\xi^{\mu-\beta} \dfrac{O(1)}{\rho_\mu}, & \text{if } \beta \neq 0, \\ \Psi(z)\xi^\mu \log \xi \dfrac{O(1)}{\rho_\mu}, & \text{if } \beta = 0. \end{cases}$$

The associated result

$$(39') \qquad u'_{2m,j}(z) = y'_{2m,j}(z) + \begin{cases} \dfrac{\xi^{-\mu-\beta}}{\Psi(z)} \dfrac{\rho^{2\mu}O(1)}{\rho_\mu}, & \text{if } \beta \neq 0, \\ \dfrac{\xi^{-\mu} \log \xi}{\Psi(z)} \dfrac{\rho^{2\mu}O(1)}{\rho_\mu}, & \text{if } \beta = 0, \end{cases}$$

is obtainable by the substitution of (39) into the derived relation (20′).

The formulas (39), (39′) give the descriptions of the solutions $u_{k,j}(z)$ for values of z such that $|\xi| \leqq N$. It only remains, therefore, to consider the forms of these solutions when $|\xi| > N$ and ξ is not in one of the specially associated sub-regions (13) indicated in the concluding formulas of §7. In this connection the following considerations may be made.

Every value ξ is included in some sub-region (13), and hence any given ξ may be said to lie in $\Xi^{(h)}$ since this amounts merely to a specification of the index h. The associated solutions $u_{h,j}(z)$ are then of the forms which have been deduced above. In terms of them the solutions of any arbitrarily chosen pair, say $u_{k,j}(z)$, may be expressed linearly, i.e.,

$$(40) \qquad u_{k,j}(z) \equiv C_{1,j}u_{h,1}(z) + C_{2,j}u_{h,2}(z),$$

with coefficients $C_{i,j}$ independent of z. The identities between the corresponding solutions of the related equation may be written

$$(41) \qquad y_{k,j}(z) \equiv C'_{1,j}y_{h,1}(z) + C'_{2,j}y_{h,2}(z).$$

Now when $|\xi| \leqq N$ the solutions involved in (40) are given by the formulas (39), and hence (40) is also expressible in the form

$$y_{k,j}(z) + \Psi(z)\xi^{\mu-\beta}\frac{O(1)}{\rho_\mu} \equiv C_{1,j}y_{h,1}(z) + C_{2,j}y_{h,2}(z).$$

On subtracting (41) from this it is accordingly found that

$$(C_{1,j} - C'_{1,j})y_{h,1}(z) + (C_{2,j} - C'_{2,j})y_{h,2}(z) \equiv \Psi(z)\xi^{\mu-\beta}\frac{O(1)}{\rho_\mu},$$

a relation which in virtue of the forms (12) implies that

$$C_{i,j} = C'_{i,j} + \frac{O(1)}{\rho_\mu}, \qquad i, j = 1, 2.$$

On inserting these evaluations in (40), however, and allowing ξ to take large values in $\Xi^{(h)}$ the comparison of the right-hand members of (40) and (41) leads to the conclusion that

$$(42)\quad u_{k,j}(z) = y_{k,j}(z) + \Psi(z)\xi^{\mu-1/2}\left\{\frac{e^{i\xi}O(1) + e^{-i\xi}O(1)}{\rho_\mu}\right\}, \quad \text{when } |\xi| > N.$$

Since the index h is not in evidence, this formula is valid for an arbitrary location of ξ in the region $|\xi| > N$. Finally, if the forms (15) are inserted for the solutions $y_{k,j}$ the results are the following:

$$(43)\quad \begin{aligned} &u_{2m,1}(z) \equiv u_{2m+1,1}(z), \qquad u_{2m-1,2}(z) \equiv u_{2m,2}(z), \\ &u_{2m,j}(z) = \Psi(z)\xi^{\mu-1/2}\left\{e^{i\xi}\left[c_{j,1}^{(m)} + O\left(\frac{1}{\xi}\right) + O\left(\frac{1}{\rho_\mu}\right)\right]\right. \\ &\qquad\qquad \left. + e^{-i\xi}\left[c_{j,2}^{(m)} + O\left(\frac{1}{\xi}\right) + O\left(\frac{1}{\rho_\mu}\right)\right]\right\}, \end{aligned}$$

with coefficients dependent upon the location of ξ as given in (16). These formulas include the formulas (37). However, when they are applicable the latter are more precise, since in (43) the vanishing of the entire coefficient of an exponential cannot be directly inferred from the vanishing of the constant $c_{j,1}^{(m)}$ involved, and the identification of a solution as of the sub-dominant type may thereby be made impossible.

THEOREM 2. *Under the hypotheses* (i) *to* (vi) *the differential equation* (2) *admits of fundamental pairs of solutions* $u_{k,j}(z)$, $j=1, 2$; $k=0, \pm 1, \pm 2, \cdots$, *which are of the forms* (39), (39′) *for values of* z *such that* $|\xi| \leqq N$, *and for values of* z *such that* $|\xi| > N$ *are generally of the forms* (43) *with the coefficients* (16), *and more specifically of the forms* (37) *in the sub-regions for which the latter are indicated to be valid.*

UNIVERSITY OF WISCONSIN,
MADISON, WIS.

A PRIORI LIMITATIONS FOR SOLUTIONS OF MONGE-AMPÈRE EQUATIONS*

BY

HANS LEWY

Introduction. As has been well known since the fundamental papers of S. Bernstein† on the generalization of Dirichlet's principle, one of the main points in proofs of existence for elliptic equations, and the most difficult one, is to obtain suitable estimates for solutions and their derivatives *under the assumption* of their existence, the so-called a priori limitations. Since the purpose of this paper is merely to establish such limitations in the case of analytic elliptic Monge-Ampère equations, we do not intend to repeat here once more the classical reasoning of S. Bernstein.

The method used in this paper will be found different from those generally adopted for similar purposes in the case of quasi-linear equations. It seems to the author that the adaptation of a Monge-Ampère equation to a scheme permitting the application of the sharp a priori limitations for linear equations‡ is efficient as long as there exist estimates for the third derivatives, but is no longer available, if, as in the present paper, we try to *establish* estimates for the third derivatives assuming the knowledge of bounds for the absolute values of the solution and its derivatives up to the second order.

Our method consists in an analytic continuation of the solution into a complex domain similar to that which we introduced in a former paper§ in order to prove the uniqueness of Cauchy's problem. In the case of a Monge-Ampère equation we can, however, use a simpler system of characteristic equations than in the most general non-linear case of analytic elliptic equations. After this has been done a study of the behavior of the analytic continuation is undertaken in order to prove that it has a certain property of "Schlichtheit." This is the principal feature which makes the application of Cauchy's integral formula possible. This Schlichtheit is proved by some simple topological considerations.

* Presented to the Society, September 5, 1934; received by the editors October 1, 1934.

† S. Bernstein, *Sur la généralisation du problème de Dirichlet*, Mathematische Annalen, vol. 62 (1906), p. 253, and vol. 69 (1910), p. 82.

‡ See for example Julius Schauder, *Über lineare elliptische Differentialgleichungen zweiter Ordnung*, Mathematische Zeitschrift, vol. 38 (1934), p. 257.

§ Hans Lewy, *Eindeutigkeit der Lösung des Anfangsproblems einer elliptischen Differentialgleichung*, Mathematische Annalen, vol. 104 (1931), p. 325.

In a subsequent paper we intend to give an application of our present results to an important existence problem of differential geometry in the large.

1. We start with a discussion of the analytic continuation of solutions mentioned in the Introduction. Let A, B, C, D, E be analytic functions of the five complex arguments u, v, x, p, q in a complex neighborhood N of a real point $T_0(u_0, v_0, x_0, p_0, q_0)$, $p_0^2+q_0^2 \leqq 1$, which assume real values for real values of the arguments. Suppose moreover that at the point T_0 the expression

$$\Delta \equiv 4(AC + DE) - B^2 = 2\alpha^2, \ \alpha > 0,$$

so that

$$4(AC + DE) \geqq 2\alpha^2.$$

Now let $M>0$ be a given number. Then there exists a sufficiently small subneighborhood N_ϵ of T_0 of the form

$$| u - u_0 | \leqq \epsilon, \ | v - v_0 | \leqq \epsilon, \ | x - x_0 - p_0(u - u_0) - q_0(v - v_0) | \leqq 4\epsilon^2 M,$$
$$| p - p_0 | \leqq 2\epsilon M, \ | q - q_0 | \leqq 2\epsilon M, \ \epsilon > 0,$$

and a constant $K>0$, such that at every point T of N_ϵ,

$$| \Delta | > \alpha^2, \qquad | 4(AC + DE) | > \alpha^2, \tag{1}$$

and that all the functions A, B, C, D, E are analytic in N_ϵ and remain, together with their first partial derivatives, less than K in absolute value. We omit the obvious proof and state that the numbers α, ϵ, K are functionals of A, B, C, D, E and functions of the number M. In N_ϵ the two roots of the quadratic equation in λ

$$\lambda^2 - \lambda B + (AC + DE) = 0 \tag{2}$$

are distinct and given by

$$2\lambda_1 = B + i\Delta^{1/2} \text{ and } 2\lambda_2 = B - i\Delta^{1/2}.$$

We now take for x an analytic and real-valued function of two real variables u, v, defined in the square

$$S_\epsilon : \qquad | u - u_0 | \leqq \epsilon, \qquad | v - v_0 | \leqq \epsilon,$$

and denote, as usual, by p, q, r, s, t its first and second partial derivatives. We suppose

$$x(u_0, v_0) = x_0, \qquad p(u_0, v_0) = p_0, \qquad q(u_0, v_0) = q_0,$$

and assume that in the square S_ϵ

$$| r |, \ | s |, \ | t | \leqq M \tag{1.1}$$

and that throughout S_ϵ the Monge-Ampère equation holds:

$$Ar + Bs + Ct + D(rt - s^2) = E. \tag{3}$$

Condition (1.1) implies that, for every point of S_ϵ, (u, v, x, p, q) lies in N_ϵ. We take two points of S_ϵ, P and Q, and map the line segment PQ affinely onto the line segment from the point $(-1, 1)$ to the point $(1, -1)$ of the line $\gamma+\delta=0$ of the γ, δ-plane. Thus to every point π of the latter segment there belongs a pair of functions $u(\pi)$, $v(\pi)$, which serve to define three more functions of π,

$$x(\pi) = x(u(\pi), v(\pi)), \qquad p(\pi) = p(u(\pi), v(\pi)), \qquad q(\pi) = q(u(\pi), v(\pi)).$$

With that segment as initial curve and $u(\pi)$, $v(\pi)$, $x(\pi)$, $p(\pi)$, $q(\pi)$ as initial values we set up the following hyperbolic system (of characteristic equations of (3)):

$$x_\gamma - pu_\gamma - qv_\gamma = 0, \tag{4.1}$$

$$\lambda_1 u_\gamma - Av_\gamma - Dq_\gamma = 0, \tag{4.2}$$

$$\lambda_2 u_\delta - Av_\delta - Dq_\delta = 0, \tag{4.3}$$

$$- Ev_\gamma + Cq_\gamma + \lambda_1 p_\gamma = 0, \tag{4.4}$$

$$- Ev_\delta + Cq_\delta + \lambda_2 p_\delta = 0. \tag{4.5}$$

Concerning this problem we prove the following

THEOREM. *There exists a positive number $\epsilon_1<\epsilon$, which depends only on ϵ, α, K, M, such that*
(i) *a solution $u(\gamma, \delta)$, $v(\gamma, \delta)$, $\cdots$, $q(\gamma, \delta)$ of* (4.1–4.5) *exists throughout the square $|\gamma| \leqq 1$, $|\delta| \leqq 1$;*
(ii) *for every point of this square $(u, \cdots, q)$ belongs to N_ϵ;*
(iii) *the solution has continuous derivatives with respect to γ and δ and depends analytically on the coordinates of the points P and Q, used in the determination of the initial values;*
(iv) *in the square $|\gamma| \leqq 1$, $|\delta| \leqq 1$ the first derivatives of the functions u, v, x, p, q with respect to γ and δ remain bounded by $2\tau z$, $z=\overline{PQ}$, and moreover*

$$|u_\gamma(\gamma, \delta) - u_\gamma(\gamma, -\gamma)| \leqq \beta z, \tag{5}$$

provided P and Q lie both in the square

$$S_{\epsilon_1}: \qquad |u - u_0| \leqq \epsilon_1, \qquad |v - v_0| \leqq \epsilon_1;$$

the numbers τ, β are functions of α, K, M, ϵ, to be specified later.

We first derive certain bounds for our functions $x(u, v)$, $p(u, v)$, $q(u, v)$, $r(u, v)$, $s(u, v)$, $t(u, v)$ for real u, v in S_ϵ. In view of (3),

$$\Delta = 4(A + Dt)(C + Dr) - (B - 2Ds)^2.$$

Hence, by (1),

$$4(A + Dt)(C + Dr) > \alpha^2.$$

The factor $A+Dt$, being continuous, never vanishes and may, without loss of generality, be assumed positive. We then have

$$(6) \qquad A + Dt > \frac{\alpha^2}{4K(1 + M)}, \qquad C + Dr > \frac{\alpha^2}{4K(1 + M)},$$

$$(6') \qquad \left(\frac{\frac{B}{2} - Ds}{A + Dt}\right)^2 = \frac{-\Delta}{4(A + Dt)^2} + \frac{C + Dr}{A + Dt} < \frac{C + Dr}{A + Dt} \leqq \frac{4K^2(1 + M)^2}{\alpha^2},$$

$$(7) \qquad |\Delta| \leqq 9K^2, \qquad |\Delta|^{1/2} \leqq 3K,$$

the last inequality being valid in the whole N_ϵ. We now define a non-negative angle $\phi_1 < \pi/2$ by the equation

$$(8) \qquad \tan \phi_1 = \frac{4K(1 + M)}{\alpha}.$$

If $|\phi| \leqq \phi_1$, we have $\cos \phi \geqq \cos \phi_1 > 0$. If, on the other hand, $\pi - \phi_1 \geqq \phi \geqq \phi_1$, we deduce the following inequalities:

$$\tan \phi_1 \leqq \tan \phi, \quad \left|\cos \phi \cdot \left(Ds - \frac{B}{2}\right)\right| \leqq \frac{1}{2} \sin \phi \cdot (A + Dt),$$

$$(9) \qquad \cos \phi \cdot \left(Ds - \frac{B}{2}\right) + \sin \phi \cdot (A + Dt) \geqq \frac{1}{2} \sin \phi \cdot (A + Dt)$$

$$\geqq \frac{1}{2} \sin \phi_1 \cdot (A + Dt) \geqq \frac{\alpha^2 \sin \phi_1}{8K(1 + M)}.$$

In order to solve the above stated hyperbolic problem, we reduce it to a system of five equations of second order. Thus we shall need to know the initial values of the first derivatives of the solution whose existence we are going to establish, along the initial line $\gamma + \delta = 0$. The equation of the line in the u, v plane joining P and Q has the form

$$v \cos \phi - u \sin \phi = \text{const.},$$

where ϕ is the angle of the direction from P to Q with the positive u axis. On the other hand let us denote by $(')$ the operation of differentiation with respect to the arc length of the line $\gamma + \delta = 0$. We evidently have for any function $f(\gamma, \delta)$, having continuous first derivatives,

$$f' = \frac{1}{2^{1/2}}\left(\frac{\partial}{\partial\gamma} - \frac{\partial}{\partial\delta}\right)f.$$

Thus we find

$$2^{1/2}u' = u_\gamma - u_\delta = \frac{z}{2}\cos\phi, \qquad 2^{1/2}v' = v_\gamma - v_\delta = \frac{z}{2}\sin\phi,$$

$$2^{1/2}p' = \frac{z}{2}(r\cos\phi + s\sin\phi), \quad 2^{1/2}q' = \frac{z}{2}(s\cos\phi + t\sin\phi),$$

$$2^{1/2}x' = p(u_\gamma - u_\delta) + q(v_\gamma - v_\delta).$$

This together with the equations (4.1–4.5) yields the following values for the initial values of u_γ, u_δ, v_γ, v_δ, account being taken of the non-vanishing of the determinant

$$(10)\qquad \begin{vmatrix} 1 & -p & -q & 0 & 0 \\ 0 & \lambda_1 & -A & -D & 0 \\ 0 & \lambda_2 & -A & -D & 0 \\ 0 & 0 & -E & C & \lambda_1 \\ 0 & 0 & -E & C & \lambda_2 \end{vmatrix} = (\lambda_1 - \lambda_2)^2(AC + DE)$$

$$= -\Delta(AC + DE);$$

$$(11)\qquad \begin{aligned} u_\gamma &= \frac{z}{2(\lambda_1 - \lambda_2)}[-(\lambda_2 - Ds)\cos\phi + (A + Dt)\sin\phi], \\ u_\delta &= \frac{z}{2(\lambda_1 - \lambda_2)}[-(\lambda_1 - Ds)\cos\phi + (A + Dt)\sin\phi], \\ v_\gamma &= \frac{z}{2(\lambda_1 - \lambda_2)}[-(Dr + C)\cos\phi + (\lambda_1 - Ds)\sin\phi], \\ v_\delta &= \frac{z}{2(\lambda_1 - \lambda_2)}[-(Dr + C)\cos\phi + (\lambda_2 - Ds)\sin\phi], \end{aligned}$$

$$x_\gamma = pu_\gamma + qv_\gamma, \qquad x_\delta = pu_\delta + qv_\delta,$$

$$p_\gamma = ru_\gamma + sv_\gamma, \qquad p_\delta = ru_\delta + sv_\delta,$$

$$q_\gamma = su_\gamma + tv_\gamma, \qquad q_\delta = su_\gamma + tv_\delta.$$

This leads to an important set of inequalities for the real and imaginary parts of u_γ on the initial line $\gamma+\delta=0$. If $|\phi| \leqq \phi_1$, we have

$$(12_1)\qquad \Re(u_\gamma) = \frac{z}{4}\cos\phi \geqq \frac{z}{4}\cos\phi_1.$$

Similarly for $\pi-\phi_1 \leqq \phi \leqq \pi+\phi_1$,

$$(12_2) \qquad \Re(u_\gamma) = \frac{z}{4}\cos\phi \leqq -\frac{z}{4}\cos\phi_1.$$

If, on the other hand, $\phi_1 \leqq \phi \leqq \pi-\phi_1$, we find, on account of (9) and (7),

$$(12_3) \qquad -\Im(u_\gamma) \geqq \frac{\alpha^2 \sin\phi_1}{96K^2(1+M)}\cdot z,$$

and, for $\pi+\phi_1 \leqq \phi \leqq 2\pi-\phi_1$,

$$(12_4) \qquad -\Im(u_\gamma) \leqq \frac{-\alpha^2 \sin\phi_1}{96K^2(1+M)}\cdot z.$$

This may be expressed in the following form: the entire range of ϕ splits up into four sectors such that within the first sector $\Re(u_\gamma)$ is bounded below by $2\beta z$, within the third sector $\Re(u_\gamma)$ is bounded above by $-2\beta z$, within the second sector $-\Im(u_\gamma)$ is bounded below by $2\beta z$ and within the fourth sector $-\Im(u_\gamma)$ is bounded above by $-2\beta z$. It is sufficient to choose for $2\beta>0$ the smaller one of the two numbers $\frac{1}{4}\cos\phi_1$ and $\alpha^2\sin\phi_1/(96K^2(1+M))$ which both depend only on α, K, M. Hence β depends only on α, K and M.

We notice furthermore that our formulas (11) show the existence of a positive number τ, depending only on α, K, M, ϵ, such that, on $\gamma+\delta=0$,

$$(13) \qquad |u_\gamma| < \tau z, \quad |u_\delta| < \tau z, \quad \cdots, \quad |q_\delta| < \tau z.$$

We now indicate briefly a method† of solving the system (4.1–4.5). Differentiate (4.1), (4.2), (4.4) with respect to δ and (4.3) and (4.5) with respect to γ. We obtain five equations containing second derivatives of the unknown functions only of the type $\partial^2/\partial\gamma\partial\delta$. We solve with respect to these, which is possible since $\Delta\neq 0$. The equations so obtained have the form

$$(14) \qquad \Delta u_{\gamma\delta} = \text{quadratic form in } u_\gamma, u_\delta, \cdots, q_\gamma, q_\delta; \cdots,$$

with coefficients which are polynomials in A, B, C, D, E and their partial derivatives with respect to u, v, x, p, q, in other words with coefficients limited in absolute value by a suitable polynomial $g(K)>0$ as long as u, v, x, p, q remain within N_ϵ. Under the same condition $|\Delta|>\alpha^2$. We now apply successive approximations to system (14) with initial values as determined above by two points P and Q of S_{ϵ_1} where ϵ_1 satisfies the following inequalities:

$$(15) \qquad \begin{gathered} \epsilon_1 < \epsilon, \qquad 800g(K)\tau\cdot 2^{3/2}\epsilon_1 < \alpha^2, \qquad \epsilon_1(1+8\cdot 2^{1/2}\tau) < \epsilon, \\ \epsilon_1(2M+8\cdot 2^{1/2}\tau) < 2M\epsilon, \qquad \epsilon_1(4M\epsilon+24\cdot 2^{1/2}\tau) < 4M\epsilon^2, \\ 800g(K)\tau^2\cdot 2^{3/2}\epsilon_1 < \alpha^2\beta. \end{gathered}$$

† Cf. Hadamard, *Leçons sur le Problème de Cauchy*, Paris, Hermann, 1932, pp. 488–501.

Postponing the definition of the first approximation, let us suppose that any one of the successive approximations exists in $|\gamma| \leqq 1$, $|\delta| \leqq 1$, and that (u, v, x, p, q) is in N_ϵ, while the partial derivatives with respect to γ or δ are in absolute value $\leqq 2\tau z$. Then for the derivatives of the next approximation $u^*, \cdots, q^*$ which certainly exists in $|\gamma| \leqq 1$, $|\delta| \leqq 1$, we find from (14)

$$\left| \overset{*}{u}_{\gamma\delta} \right| \leqq \frac{g(K)}{\alpha^2} \cdot 400\tau^2 z^2.$$

On integrating this and taking into account (13) and the relation $\overline{PQ} = z \leqq 2^{3/2}\epsilon_1$, we have

$$\begin{aligned} &| u_\gamma^* | \leqq \tau z + \frac{800 g(K)}{\alpha^2}\tau^2 z^2 \leqq \tau z\left(1 + \frac{800 g(K)}{\alpha^2}\tau z\right) \leqq 2\tau z, \\ &| u_\delta^* | \leqq 2\tau z, \cdots, | q_\delta^* | \leqq 2\tau z \qquad \text{for } |\gamma| \leqq 1, \ |\delta| \leqq 1. \end{aligned} \tag{16}$$

On the other hand by (1.1) we have for the initial values of the u, v, x, p, q

$$\begin{aligned} &| u - u_0 | \leqq \epsilon_1, \ | v - v_0 | \leqq \epsilon_1, \ | p - p_0 | \leqq 2M\epsilon_1, \ | q - q_0 | \leqq 2M\epsilon_1, \\ &| x - x_0 - p_0(u - u_0) - q_0(v - v_0) | \leqq 4M\epsilon_1^2. \end{aligned}$$

Hence on integrating (16) and in view of (15) and $|p_0| \leqq 1$, $|q_0| \leqq 1$, we have for $|\gamma| \leqq 1$, $|\delta| \leqq 1$,

$$\begin{aligned} &| p^* - p_0 | \leqq 2M\epsilon_1 + 8\tau \cdot 2^{1/2}\epsilon_1 \leqq 2M\epsilon, \\ &| q^* - q_0 | \leqq 2M\epsilon, \\ &| u^* - u_0 | \leqq \epsilon_1 + 8\tau \cdot 2^{1/2}\epsilon_1 \leqq \epsilon, \qquad | v^* - v_0 | \leqq \epsilon, \\ &| x^* - x_0 - p_0(u^* - u_0) - q_0(v^* - v_0) | \leqq 4M\epsilon_1^2 + 24 \cdot 2^{1/2}\tau\epsilon_1 \leqq 4M\epsilon^2. \end{aligned} \tag{17}$$

This shows that all the functions u^*, v^*, x^*, p^*, q^* remain within N_ϵ, provided we define a suitable first approximation. We do this by setting as a first approximation a solution of the system

$$u_{\gamma\delta} = 0, \ v_{\gamma\delta} = 0, \cdots, \ q_{\gamma\delta} = 0,$$

with the given initial data.

Since for this solution the values of $u_\gamma, \cdots, q_\delta$ at any point γ, δ coincide with some initial value, the above inequalities (16), (17) hold for the first approximation and thus for all of them.

It may now be shown that the successive approximations converge uniformly together with their first derivatives to limit functions and we get a solution of the system (4) assuming the given initial values. This solution is the only one having continuous first derivatives and can be differentiated

with respect to γ and δ as well as with respect to any parameter as many times as the initial data themselves.†

Obviously the solution satisfies the same inequalities (16) and (17) that are satisfied by all the approximations.

Moreover, by (14),

$$\left| u_\gamma(\gamma, \delta) - u_\gamma(\gamma, -\gamma) \right| = \left| \int_\delta^{-\gamma} \frac{\partial^2 u}{\partial\gamma\partial\delta} d\delta \right| \leqq 800 \frac{g(K)}{\alpha^2} \tau^2 z^2 \leqq \beta z.$$

This is the desired inequality (5). Thus all the statements of our theorem on page 419 have been proved except the analytical dependence on the coordinates of P and Q.

This last property results from the fact that our initial data are analytic functions of these four coordinates. Let us denote the coordinates of P by a_1 and b_1, those of Q by a_2 and b_2. If we give a_1, b_1, a_2, b_2 complex values, but *very near* to those real values for which we made the above construction, we get new initial data which differ as little as we please from the former ones; we may assume that the inequalities (13) remain valid for the new initial data. If now we construct the solution of the modified initial problem, we get again a solution for the whole square $|\gamma| \leqq 1$, $|\delta| \leqq 1$ and this solution stays as near as we please to the former solution on account of the continuous dependence on the initial data. We thus get a solution depending on four complex parameters and the two real variables γ, δ. Denote by $\eta = \eta_1 + i\eta_2$ any one of the complex parameters. The initial values, being analytic functions of η, satisfy the Cauchy-Riemann equations which evidently can be written in the following form using the abbreviation $\nabla = \partial/\partial\eta_1 + i\partial/\partial\eta_2$:

$$\nabla u = 0, \quad \nabla v = 0, \quad \cdots, \quad \nabla q = 0.$$

On the other hand, on operating with ∇ on the system (4.1–4.5) we get

$$\frac{\partial}{\partial\gamma}(\nabla x) - p\frac{\partial}{\partial\gamma}(\nabla u) - q\frac{\partial}{\partial\gamma}(\nabla v) = u_\gamma \nabla p + v_\gamma \nabla q,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$-E\frac{\partial}{\partial\delta}(\nabla v) + C\frac{\partial}{\partial\delta}(\nabla q) + \lambda_2\frac{\partial}{\partial\delta}(\nabla p) = v_\gamma \frac{\partial E}{\partial u}\nabla u + \cdots,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

i.e., a hyperbolic system, linear and homogeneous, with a non-vanishing "determinant" (10) for the unknowns $\nabla x, \cdots, \nabla q$. Since a solution of this system is $\nabla x = \cdots = \nabla q = 0$ and the solution is uniquely determined by initial values, we conclude

† For the proof of all these statements, see Hadamard, loc. cit.

$$\nabla x = \nabla u = \cdots = \nabla q = 0.$$

This implies the desired analytic dependence on a_1, b_1, a_2, b_2 in the square $|\gamma| \leqq 1$, $|\delta| \leqq 1$, expressed by means of the Cauchy-Riemann equations.

2. We return to real values of a_1, b_1, a_2, b_2. The manifold of solutions seems to depend on these four variables and the two variables γ, δ. It is possible, however, to reduce the dependence to that of a_1, b_1, a_2, b_2 only. If we substitute for γ a variable $\gamma' = k\gamma + l$ and for δ a variable $\delta' = k'\delta + l'$ with $kk' \neq 0$, a solution of the system (4) becomes a solution of the same system in γ', δ' because of the homogeneity of (4) with respect to the derivatives $\partial/\partial\gamma$, $\partial/\partial\delta$. Now let P', Q' be two points of, and in the same order as, PQ. The range of the initial values of the problem (4) determined by P' and Q' may be transformed by the above substitution into a part of the range of the initial values of the PQ problem. The same must then be true for the range of the corresponding solutions of the $P'Q'$ problem and of the PQ problem. If therefore we agree to consider only the point $\gamma = 1$, $\delta = 1$ of the PQ problem and to write the solution at this point as a function of a_1, b_1, a_2, b_2 only, the whole square will consist of points at which the solution is the same function with different values of the argument a_1, b_1, a_2, b_2. In particular, the line $\delta = 1$ corresponds to fixed (a_1, b_1) while (a_2, b_2) varies along a line through (a_1, b_1). This shows that we have on $\delta = 1$ the following rule: The differentiation $\partial/\partial\gamma$ of our solution, considered as function of γ, δ and of the four parameters a_1, b_1, a_2, b_2, reduces to the operation

$$\frac{z}{2}\left(\cos\phi \frac{\partial}{\partial a_2} + \sin\phi \frac{\partial}{\partial b_2}\right)$$

if we write our solutions as functions of a_1, b_1, a_2, b_2 only. Notice that the solution is defined for $|a_1 - u_0| \leqq \epsilon_1$, $|b_1 - v_0| \leqq \epsilon_1$, $|a_2 - u_0| \leqq \epsilon_1$, $|b_2 - v_0| \leqq \epsilon_1$ and that the originally given functions in the real u, v-plane now appear as functions of the four arguments for coincident P and Q, i.e., for $a_1 = a_2$, $b_1 = b_2$.

3. We now set up another hyperbolic system for two functions $u(\gamma, \delta)$ and $v(\gamma, \delta)$ to be determined for $|\gamma| \leqq 1$, $|\delta| \leqq 1$. The initial values are given on $\gamma + \delta = 0$ in the same way as on page 419, but the equations are

$$(\lambda_1 - Ds)u_\gamma - (A + Dt)v_\gamma = 0, \tag{18.1}$$

$$(\lambda_2 - Ds)u_\delta - (A + Dt)v_\delta = 0. \tag{18.2}$$

Here the quantities x, p, q, r, s, t are to be considered as analytic functions of u and v, determined by the originally presented solution of the Monge-Ampère equation. On writing

(19) $$\rho = \frac{\lambda_1 - Ds}{A + Dt}, \qquad \bar{\rho} = \frac{\lambda_2 - Ds}{A + Dt},$$

this becomes

(18.3) $$v_\gamma - \rho u_\gamma = 0,$$

(18.4) $$v_\delta - \bar{\rho} u_\delta = 0.$$

We conclude

$$\lambda_1 u_\gamma - A v_\gamma - D q_\gamma = 0, \quad \lambda_2 u_\delta - A v_\delta - D q_\delta = 0, \quad - E v_\gamma + C q_\gamma + \lambda_1 p_\gamma = 0,$$

since, in view of (2) and (3),

$$(\lambda_1 r + Cs)(A + Dt) + (\lambda_1 s + Ct - E)(\lambda_1 - Ds) = 0.$$

Similarly we obtain the relation

$$- E v_\delta + C q_\delta + \lambda_2 p_\delta = 0.$$

We also mention the identity

$$x_\gamma - p u_\gamma - q v_\gamma = 0.$$

Thus the introduction of the new initial problem leads to a set of functions $u(\gamma, \delta), \cdots, q(\gamma, \delta)$ which turns out to be a solution of (4.1–4.5) with the same initial conditions. In view of the uniqueness theorem the two solutions must coincide wherever they exist simultaneously, which is true for P sufficiently near Q.

Equations (18.3) and (18.4) admit of a simple interpretation. For instance, if γ varies alone, the differentials of u and v are connected by the *ordinary* differential equation

$$dv - \rho du = 0.$$

Returning to the variables a_1, b_1, a_2, b_2, we may say that if (a_1, b_1) is fixed while (a_2, b_2) varies along any line through (a_1, b_1), then (u, v) moves on a certain two-dimensional surface M_1 of the four-dimensional complex u, v-space, called characteristic "megaline."* For (a_2, b_2) sufficiently near (a_1, b_1) the equation of M_1 has the form

$$v = \text{analytic function of } u.$$

Thus, for Q sufficiently near P, we may write the corresponding Cauchy-Riemann equations in the form

(20) $$v_{a_2} u_{b_2} - v_{b_2} u_{a_2} = 0.$$

* Hadamard, loc. cit., p. 512.

But in view of the above stated analytical dependence on (a_1, b_1, a_2, b_2) this relation holds on the whole range of the values of (a_1, b_1, a_2, b_2). Similarly we find

$$v_{a_1}u_{b_1} - v_{b_1}u_{a_1} = 0. \tag{21}$$

From (20), (21), (18.3), (18.4) we derive a fact of fundamental importance for our discussion, namely that the *quotient* $u_{a_2}:u_{b_2}$ *is independent of* a_1, b_1.

Indeed, since on M_1 we have $0 = dv - \rho du$, the relation

$$v_{a_2} - \rho u_{a_2} = 0 \tag{22.1}$$

is satisfied on M_1 and holds for any M_1, or, in other words, for any a_1, b_1. Similarly

$$v_{b_2} - \rho u_{b_2} = 0, \tag{22.2}$$

$$v_{a_1} - \bar{\rho} u_{a_1} = 0. \tag{23.1}$$

Now differentiate (23.1) with respect to a_2 and b_2, (22.1) with respect to a_1 and (22.2) with respect to a_1. This gives

$$v_{a_1a_2} - \bar{\rho}u_{a_1a_2} - \bar{\rho}_u u_{a_1}u_{a_2} - \bar{\rho}_v u_{a_1}v_{a_2} = 0,$$
$$v_{a_1b_2} - \bar{\rho}u_{a_1b_2} - \bar{\rho}_u u_{a_1}u_{b_2} - \bar{\rho}_v u_{a_1}v_{b_2} = 0,$$

$$v_{a_1a_2} - \rho u_{a_1a_2} - \rho_u u_{a_2}u_{a_1} - \rho_v u_{a_2}v_{a_1} = 0,$$
$$v_{a_1b_2} - \rho u_{a_1b_2} - \rho_u u_{b_2}u_{a_1} - \rho_v u_{b_2}v_{a_1} = 0.$$

By (20)

$$v_{a_1a_2}u_{b_2} - v_{a_1b_2}u_{a_2} - \bar{\rho}(u_{a_1a_2}u_{b_2} - u_{a_1b_2}u_{a_2}) = 0,$$
$$v_{a_1a_2}u_{b_2} - v_{a_1b_2}u_{a_2} - \rho(u_{a_1a_2}u_{b_2} - u_{a_1b_2}u_{a_2}) = 0.$$

$\rho \neq \bar{\rho}$ gives

$$u_{a_1a_2}u_{b_2} - u_{a_1b_2}u_{a_2} = 0,$$

which shows that $u_{a_2}:u_{b_2}$ does not depend on a_1.

In an analogous manner we find that $u_{a_2}:u_{b_2}$ does not depend on b_1. Hence, in order to compute the value of $u_{a_2}:u_{b_2}$ for any (a_1, b_1, a_2, b_2) we may compute it for $a_1 = a_2$, $b_1 = b_2$, i.e., for (a_2, b_2, a_2, b_2). In view of the known values (11) for the initial values of u_γ and our remark on page 425 we get

$$\begin{aligned} \frac{2}{z} u_\gamma &= \cos\phi \cdot u_{a_2} + \sin\phi \cdot u_{b_2} \\ &= \frac{1}{\lambda_1 - \lambda_2}[-(\lambda_2 - Ds)\cos\phi + (A + Dt)\sin\phi]. \end{aligned} \tag{24}$$

Taking $\phi=0$ and $\phi=\pi/2$ gives finally

$$u_{a_2}:u_{b_2} = -\frac{\lambda_2 - Ds}{A + Dt}. \tag{25}$$

We notice that this expression has a positive imaginary part. Since x, p, q are analytic functions of u and v and thus of u alone on M_1, we have on M_1 the Cauchy-Riemann equations

$$\begin{aligned} x_{a_2}u_{b_2} - x_{b_2}u_{a_2} &= 0, \\ p_{a_2}u_{b_2} - p_{b_2}u_{a_2} &= 0, \\ q_{a_2}u_{b_2} - q_{b_2}u_{a_2} &= 0, \end{aligned} \tag{23}$$

and these relations again hold for all a_1, b_1, a_2, b_2 in question.

4. We now proceed to establish a fact which forms the salient point of the present investigation.

THEOREM. *The function $u(u_0, v_0, a_2, b_2)$ is "schlicht" for $|u-u_0|<2\beta\epsilon_1$, or, in other words, the equation*

$$u(u_0, v_0, a_2, b_2) = U$$

has one and only one solution (a_2, b_2) in $|a_2-u_0|\leqq\epsilon_1$, $|b_2-v_0|\leqq\epsilon_1$ provided $|U-u_0|<2\beta\epsilon_1$.

Let us write (5) in the form

$$u_\gamma(\gamma, \delta) = u_\gamma(\gamma, -\gamma) + h, \quad |h| \leqq \beta z.$$

The value $u_\gamma(\gamma, -\gamma)$ satisfies the inequalities (12_1–12_4). Hence, for $|\gamma|\leqq 1$, $|\delta|\leqq 1$,

$$\begin{aligned} \Re(u_\gamma) &\geqq \beta z && \text{for} \quad |\phi| \leqq \phi_1, \\ \Re(u_\gamma) &\leqq -\beta z && \text{for} \quad \pi - \phi_1 \leqq \phi \leqq \pi + \phi_1, \\ -\Im(u_\gamma) &\geqq \beta z && \text{for} \quad \phi_1 \leqq \phi \leqq \pi - \phi_1, \\ -\Im(u_\gamma) &\leqq -\beta z && \text{for} \quad \pi + \phi_1 \leqq \phi \leqq 2\pi - \phi_1. \end{aligned} \tag{26}$$

An integration from $\gamma=-1$ to $\gamma=+1$ along $\delta=1$ yields

$$\begin{aligned} \Re[u(1, 1) - u_0] &\geqq 2\beta z && \text{for} \quad |\phi| \leqq \phi_1, \\ \Re[u(1, 1) - u_0] &\leqq -2\beta z && \text{for} \quad \pi - \phi_1 \leqq \phi \leqq \pi + \phi_1, \\ -\Im[u(1, 1) - u_0] &\geqq 2\beta z && \text{for} \quad \phi_1 \leqq \phi \leqq \pi - \phi_1, \\ -\Im[u(1, 1) - u_0] &\leqq -2\beta z && \text{for} \quad \pi + \phi_1 \leqq \phi \leqq 2\pi - \phi_1. \end{aligned} \tag{27}$$

Let us take for P the point (u_0, v_0) and let (a_2, b_2) vary in the square R: $|a_2-u_0|\leqq\epsilon_1$, $|b_2-v_0|\leqq\epsilon_1$. Imagine the complex conjugate quantity $\bar{u}(u_0, v_0,$

$a_2, b_2)-u_0$ as a vector* attached to the variable point (a_2, b_2) of R. Formulas (27) show that $\bar{u}(u_0, v_0, a_2, b_2)-u_0$ forms an angle less than π with the vector

$$a_2 - u_0 + i(b_2 - v_0)$$

which is $\neq 0$, provided $(a_2-u_0)^2+(b_2-v_0)^2\neq 0$. If we follow $\bar{u}$ along any closed Jordan curve contained in R and such that (u_0, v_0) lies in its interior, the rotation of $\bar{u}-u_0$ after one circuit must equal that of $a_2-u_0+i(b_2-v_0)$, i.e., 2π. On the other hand the vector $\bar{u}-u_0$ depends analytically on a_2, b_2 and has for every subdomain of R only a finite number of singularities. (The same will be true for any vector $\bar{u}(u_0, v_0, a_2, b_2) - \bar{U}$, $\bar{U}$ being a constant.) We conclude that the point $a_2=u_0$, $b_2=v_0$ is a singularity of the field $\bar{u}(u_0, v_0, a_2, b_2)-u_0$ with index $+1$.

By (27), the expression $\bar{u}-\bar{U}$ does not vanish along the contour of R if $|U-u_0|<2\beta\epsilon_1$, since $z=\overline{PQ}\geqq\epsilon_1$ as long as Q varies along the contour of R and P is the center (u_0, v_0) of R. For reasons of continuity, the rotation (along the contour of R) of $\bar{u}-\bar{U}$ equals that of $\bar{u}-u_0$, so that we may say that the vector field $\bar{u}-\bar{U}$ has at least one singularity within R for $|U-u_0|<2\beta\epsilon_1$, and the sum of the indices of all singularities in R equals $+1$. On the other hand it follows from the lemma below that any such singularity has an essentially positive index, hence there can be at most one singularity, which is the desired result.

LEMMA. *Let, without loss of generality, the origin of an a, b-plane be a zero of a complex vector* $\mathfrak{u}(a, b)$, *analytic near the origin and such that the quotient* $\mathfrak{u}_a:\mathfrak{u}_b=\kappa$ *has a negative imaginary part at the origin*: $\mathfrak{J}(\kappa_0)<0$. *Then the singularity of the origin has a positive index.*

We prove this lemma by developing $\mathfrak{u}$ in a Taylor series in $b+\kappa_0 a$ and $b-\kappa_0 a$. Let $k>0$ be the smallest degree for which there are non-vanishing terms so that the series begins as follows:

$$\mathfrak{u} = c_1(b - \kappa_0 a)^k + c_2(b - \kappa_0 a)^{k-1}(b + \kappa_0 a) + \cdots + c_{k+1}(b + \kappa_0 a)^k + \cdots.$$

Form the expression $\mathfrak{u}_a-\kappa\mathfrak{u}_b$ and develop it into a power series in a and b; it must vanish identically. On the other hand the terms of lowest degree in $\mathfrak{u}_a-\kappa\mathfrak{u}_b$ have the form

$$- 2\kappa_0[kc_1(b - \kappa_0 a)^{k-1} + \cdots + c_k(b + \kappa_0 a)^{k-1}]$$

so that we conclude $c_1=c_2=\cdots=c_k=0$. Thus the Taylor series of $\mathfrak{u}$ begins with the term $c_{k+1}(b+\kappa_0 a)^k$, $c_{k+1}\neq 0$, and the singularity of $\mathfrak{u}$ at $a=b=0$ is that of $c_{k+1}(b+\kappa_0 a)^k$ and has the index $k>0$.

* For the notions concerning vector fields used in this paragraph, see for instance W. Fenchel, *Elementare Beweise und Anwendungen einiger Fixpunktsätze*, Matematisk Tidsskrift, (B), 1932, p. 66.

5. The theorem of the preceding section enables us to write the megaline M_1 in the form

$$v = \text{analytic function of } u$$

for u varying within the circle $\Gamma: |u-u_0| < 2\beta\epsilon_1$. Within R none of the partial derivatives of $u(u_0, v_0, a_2, b_2)$ with respect to a_2 or b_2 vanishes. For we conclude from (26) and (24) that there is at least a linear form in u_{a_2} and u_{b_2} which differs from zero. On the other hand the quotient $u_{a_2}:u_{b_2}$ has a positive imaginary part and cannot vanish, which proves that $u_{a_2}\neq 0$, $u_{b_2}\neq 0$.

Hence, the relations (20) and (23) yield the analytic dependence of v, x, p, q on u on the megaline M_1, provided u lies in Γ.

Integration of (13) gives, for $|a_2-u_0| \leqq \epsilon_1$, $|b_2-v_0| \leqq \epsilon_1$,

$$|v - v_0|, \cdots, |q - q_0| \leqq 2^{3/2}\tau\epsilon_1.$$

Let us denote by d/du the differential quotient on M_1 (u in Γ). Cauchy's integral formula yields for $u = u_0$

$$\rho = \frac{dv}{du} = \frac{1}{2\pi i}\oint_\Gamma \frac{(v - v_0)du}{(u - u_0)^2}, \quad \frac{d^n v}{du^n} = \frac{n!}{2\pi i}\oint_\Gamma \frac{(v - v_0)du}{(u - u_0)^{n+1}}, \cdots,$$

$$\frac{d^n q}{du^n} = \frac{n!}{2\pi i}\oint_\Gamma \frac{(q - q_0)du}{(u - u_0)^{n+1}};$$

$$\left|\frac{d^n v}{du^n}\right|, \cdots, \left|\frac{d^n q}{du^n}\right| \leqq \frac{n!}{2\pi}\cdot\frac{2\pi 2^{3/2}\tau\epsilon_1}{(2\beta\epsilon_1)^n} = \frac{n!2^{3/2}\tau\epsilon_1}{(2\beta\epsilon_1)^n}.$$

The differentiation d/du at the point $u = u_0$ of M_1 is connected with the partial derivatives $\partial/\partial u$ and $\partial/\partial v$ at the point u_0, v_0 by the relation

$$\frac{d}{du} = \frac{\partial}{\partial u} + \rho\frac{\partial}{\partial v}.$$

Thus the above inequalities for the successive derivatives at the point u_0 yield inequalities for the partial derivatives of the given functions $x(u, v)$, $p(u, v)$, $q(u, v)$. And anyone of these inequalities yields two estimates if we take account of the reality of these functions for real u, v.

6. As mentioned in the introduction, the most important of the a priori limitations are those for the third derivatives of x. We easily can establish them at the point u_0, v_0.

We have at u_0, v_0

$$\frac{d^2p}{du^2} = \left(\frac{\partial}{\partial u} + \rho\frac{\partial}{\partial v}\right)^2 p = \left(\frac{\partial}{\partial u} + \rho\frac{\partial}{\partial v}\right)(r + \rho s)$$

$$= r_u + 2\rho s_u + \rho^2 t_u + \frac{d^2v}{du^2}s,$$

$$\frac{d^2q}{du^2} = \left(\frac{\partial}{\partial u} + \rho\frac{\partial}{\partial v}\right)^2 q = \qquad s_u + 2\rho t_u + \rho^2 t_v + \left(\overline{\frac{d^2v}{du^2}}\right)t,$$

and the conjugate equations

$$\left(\overline{\frac{d^2p}{du^2}}\right) = \left(\frac{\partial}{\partial u} + \bar\rho\frac{\partial}{\partial v}\right)^2 p = r_u + 2\bar\rho s_u + \bar\rho^2 t_u + \left(\overline{\frac{d^2v}{du^2}}\right)s,$$

$$\left(\overline{\frac{d^2q}{du^2}}\right) = \left(\frac{\partial}{\partial u} + \bar\rho\frac{\partial}{\partial v}\right)^2 q = \qquad s_u + 2\bar\rho t_u + \bar\rho^2 t_v + \left(\overline{\frac{d^2v}{du^2}}\right)t.$$

The determinant of this system is

$$\begin{vmatrix} 1 & 2\rho & \rho^2 & 0 \\ 1 & 2\bar\rho & \bar\rho^2 & 0 \\ 0 & 1 & 2\rho & \rho^2 \\ 0 & 1 & 2\bar\rho & \bar\rho^2 \end{vmatrix} = (\rho - \bar\rho)^4 = \frac{\Delta^2}{(A + Dt)^4}$$

and is in absolute value $>\alpha^4/[K^4(1+M)^4]$ (see (1)). Moreover $|r|$, $|s|$, $|t|$ $\leqq M$ and

$$|\rho| = \left|\frac{dv}{du}\right| \leqq 2^{1/2}\tau/\beta \quad\text{and}\quad \left|\frac{d^2v}{du^2}\right|, \cdots, \left|\frac{d^2q}{du^2}\right| \leqq 2^{1/2}\tau/(\beta^2\epsilon_1).$$

Hence by solving the above linear system with respect to r_u, s_u, t_u, t_v we easily obtain upper bounds for $|r_u|$, $|s_u|$, $|t_u|$, $|t_v|$, depending on α, ϵ, K, M only. We shall not carry through the computations.

7. The estimates for the higher derivatives are not much more complicated, but for the applications it is important to know that they can be chosen in such a way that they insure a lower a priori bound for the radii of convergence of the Taylor series for x at a point u_0, v_0.

We observe that there are two megalines M_1 and M_2 through the point (u_0, v_0), conjugate to each other. The values of u, v, x, p, q at their points will be referred to by $u(u_0, v_0, a_2, b_2) \equiv u(u_0, v_0; Q), \cdots$ and $u(a_1, b_1, u_0, v_0) \equiv u(P; u_0, v_0)$ respectively. Consider the megaline passing through (u_0, v_0, Q) which consists of points $(a_1, b_1; Q)$ with variable (a_1, b_1) and fixed Q, and the megaline through $(P; u_0, v_0)$ consisting of points $(P; a_2, b_2)$ with variable

(a_2, b_2) and fixed P. The two megalines evidently intersect at (P, Q). Let ξ denote the variable u on M_1 and η the same variable on M_2. If both vary within the circle Γ of "Schlichtheit," i.e., for $|\xi-u_0|<2\beta\epsilon_1$ and $|\eta-u_0|<2\beta\epsilon_1$, they are in unique correspondence with Q and P respectively. We therefore may write $u(P, Q), \cdots$ as functions of ξ and η. We say that they are analytic functions of these complex arguments.

We have evidently

$$\xi = \xi(a_2, b_2) = \xi_1 + i\xi_2$$

and the independence of $u_{a_2}:u_{b_2}$ on a_1, b_1 shows that $u_{a_2}:u_{b_2}$ or, what is essentially the same, $u_{\xi_1}:u_{\xi_2}$ are independent on η. Hence we may compute its value by setting $\eta=u_0$. But on M_1 we have Cauchy-Riemann equations $u_{\xi_1}:u_{\xi_2}=-i$. Similarly we find the Cauchy-Riemann equations $u_{\eta_1}:u_{\eta_2}=-i$. Thus $u(\xi, \eta)$ depends analytically on ξ and η for $|\xi-u_0|<2\beta\epsilon_1$, $|\eta-u_0|<2\beta\epsilon_1$. The equations (20), (21), (23) show that for any function ω of the set u, v, x, p, q we have

$$\frac{\partial(\omega, u)}{\partial(a_2, b_2)} = 0 \quad\text{and}\quad \frac{\partial(\omega, u)}{\partial(a_1, b_1)} = 0.$$

This yields the analytic dependence of all of them on ξ and η for $|\xi-u_0|<2\beta\epsilon_1$, $|\eta-u_0|<2\beta\epsilon_1$. Moreover u, v, x, p, q admit bounds for their absolute values, depending only on ϵ, α, K, M.

On the other hand we find for $\xi=u_0$, $\eta=u_0$ the following *lower* bounds:

$$\frac{\partial u}{\partial \xi} = 1,\ \frac{\partial u}{\partial \eta} = 1,\ \left|\frac{\partial v}{\partial \xi}\right| \geqq \frac{\alpha}{2K(1+M)},\ \left|\frac{\partial v}{\partial \eta}\right| \geqq \frac{\alpha}{2K(1+M)},$$

$$\left|\frac{\partial(u, v)}{\partial(\xi, \eta)}\right| \geqq \frac{\alpha}{K(1+M)}.$$

We therefore are able to introduce u, v as independent variables* instead of ξ and η within a domain $|u-u_0|<\epsilon_2$, $|v-v_0|<\epsilon_2$, with ϵ_2 depending merely on α, ϵ, K, M. In this domain we may develop x in a power series in $u-u_0$ and $v-v_0$; it will converge there and its coefficients will have absolute bounds in terms of α, ϵ, K, M. Even the majorant series formed with these bounds instead of the derivatives of x at u_0, v_0 will converge for $|u-u_0|<\epsilon_2$, $|v-v_0|<\epsilon_2$. We state the final result in the following form:

THEOREM. *The derivatives of an analytic solution of the analytic Monge-Ampère equation* (3) *existing for* $|u-u_0|<\epsilon$, $|v-v_0|<\epsilon$ *can be developed into a power series in* $u-u_0$ *and* $v-v_0$ *whose associated radii of convergence and*

* This step is justified in §8.

whose coefficients can be bounded, the former from below, the latter from above, by certain numbers which depend only
(i) *on the bound M for the moduli of the second derivatives,*
(ii) *on the value $2\alpha^2$ of $\Delta>0$ at $u=u_0$, $v=v_0$,*
(iii) *on the bound ϵ such that in a neighborhood N_ϵ of the* 10-*dimensional space of complex x, p, q, u, v around u_0, v_0, x_0, p_0, q_0 the coefficients A, B, C, D, E remain regular, and $|\Delta|>\alpha^2$,*
(iv) *on the bound K for these coefficients and their first partial derivatives in N_ϵ, provided we have $p_0^2+q_0^2\leqq 1$. The power series in u and v, formed with the bounds of the coefficients, has the same bounds for its associated radii of convergence.*

8. We have made use of the following lemma concerning the inversion of a system of analytic functions in the neighborhood of the origin.

Lemma. *Let*

$$u = ax + by + \sum_{i+k\geqq 2} a_{ik}x^iy^k,$$

$$v = cx + dy + \sum_{i+k\geqq 2} c_{ik}x^iy^k.$$

Suppose $|ad-bc|>A>0$, $0<B_1<|a|$, $|b|$, $|c|$, $|d|<B_2$. Suppose furthermore that $\sum a_{ik}x^iy^k$ and $\sum c_{ik}x^iy^k$ converge for $|x|<\rho$, $|y|<\rho$ and that there $|u|$, $|v|<C$. Then there exists a neighborhood of the origin $|u|+|v|<h$, and a fortiori a neighborhood $|u|<h/2$, $|v|<h/2$, for which we can solve with respect to x and y, with h depending only on A, B_1, B_2, C, ρ.

Though the proof of the possibility of inversion in a sufficiently small neighborhood is well known, we were unable to find the above estimate for the neighborhood mentioned in the literature. We therefore sketch a simple method of establishing this estimate.

We have on account of the assumptions

$$|a_{ik}| < \frac{C}{\rho^{i+k}} \quad\text{and}\quad |c_{ik}| < \frac{C}{\rho^{i+k}}.$$

Hence

$$u' = \frac{du - bv}{ad - bc} = x + \sum_{i+k\geqq 2} a_{ik}'\, x^iy^k,$$

$$v' = \frac{-cu + av}{ad - bc} = y + \sum_{i+k\geqq 2} c_{ik}'\, x^iy^k,$$

and $|a_{ik}'|$, $|c_{ik}'|<C'/\rho^{i+k}$ with $C'=C'(A, B_2, C)$. For $|x|+|y|\leqq\rho/2$, we conclude

$$|u' - x| + |v' - y| < \sum_{i+k \geqq 2} \frac{2C'}{\rho^{i+k}} |x|^i |y|^k \leqq (|x| + |y|)^2 M,$$

with $M = M(C', \rho) > 0$. Upon introducing the further restriction

$$|x| + |y| \leqq \min\left\{\frac{1}{2M}, \frac{\rho}{2}\right\}$$

we get on the boundary of this domain for $|U'| + |V'| < \min\{1/(4M), \rho/4\}$,

$$|u' - x| + |v' - y| < \frac{|x| + |y|}{2},$$

$$|u' - U' - x| + |v' - V' - y| < |x| + |y|.$$

These relations show that the vector of the four-dimensional space with the components $u' - U'$, $v' - V'$ has in $|x| + |y| \leqq \min\{1/(2M), \rho/2\}$ the same sum of indices of its singularities as the vector with components x, y. This sum is $+1$. On the other hand the vector $u' - U'$, $v' - V'$, being analytic, admits only singularities with positive index. Thus there is precisely one solution, x, y, of the equations $u' = U'$, $v' = V'$ for $|U'| + |V'| < \min\{1/(4M), \rho/2\}$. This can be expressed in terms of u, v which proves our lemma.

BROWN UNIVERSITY,
PROVIDENCE, R. I.

AN ENUMERATIVE PROBLEM IN THE ARITHMETIC OF LINEAR RECURRING SERIES*

BY
MORGAN WARD

1. Let m be a fixed positive integer greater than one and let

$$\Omega_{n+k} = c_1\Omega_{n+k-1} + c_2\Omega_{n+k-2} + \cdots + c_k\Omega_n \tag{1.1}$$

be a linear difference equation of order k with rational integral coefficients $c_1, c_2, \cdots, c_k$. If

$$(U): \qquad U_0, U_1, U_2, \cdots, U_n, \cdots$$

is any sequence of rational integers satisfying (1.1), then after a certain point the sequence becomes periodic when considered modulo m. Its least period is called the characteristic number of the sequence (U) modulo m.

In a recent paper in these Transactions†, I have considered the problem of determining this characteristic number given $m, c_1, c_2, \cdots, c_k$ and the k initial values $U_0, U_1, \cdots, U_{k-1}$ of the sequence (U), and I have reduced it to certain basic problems in the theory of higher congruences.‡

In the present paper, I am concerned with the following problem which I shall similarly reduce to a problem in the theory of higher congruences:

Given any positive integer s: to find the number of distinct sequences (U) modulo m whose characteristic number is exactly equal to s.

2. I obtain here the following results.

(i) *It suffices to determine the total number of purely periodic sequences (U) modulu m whose characteristic number is at most equal to s.* (§3.)

(ii) *It suffices to confine ourselves to the case when $m = p^N$ is a power of a prime p, and when the polynomial*

$$f(x) = x^k - c_1x^{k-1} - c_2x^{k-2} - \cdots - c_k \tag{2.1}$$

of degree k associated with the difference equation (1.1) *is of the form*

$$f(x) = B(x) \equiv \{\phi(x)\}^a \pmod{p}$$

where $\phi(x)$ is irreducible modulo p and a is a positive integer. (§4.)

* Presented to the Society, December 27, 1934; received by the editors August 1, 1934.

† Vol. 35 (1933), pp. 600–628. I shall refer to this paper as Trans I.

‡ Notably, to finding the least value of n such that $A(x)(x^n-1)\equiv 0$ (modd p^N, $B(x)$) for any prime p and any assigned polynomials $A(x)$ and $B(x)$.

(iii) *The problem thus delimited is equivalent to determining the total number of distinct polynomials* $U(x)$ *of degree* $\leqq k-1$ *modulo* p^N *such that*

$$U(x)A(x) \equiv 0 \qquad (\text{modd } p^N, B(x)), \tag{2.2}$$

where

$$A(x) \equiv x^s - 1 \qquad (\text{modd } p^N, B(x)).$$

This number can be immediately written down provided that we know the elementary divisors corresponding to the prime p of the matrix $\mathcal{E}$ of the Sylvester eliminant of $A(x)$ and $B(x)$. (§5.)

In another paper in these Transactions* I have made a detailed study of the congruence (2.2) and shown that if $N \leqq \lambda$ (where p^λ is the first elementary divisor of the matrix $\mathcal{E}$ of (iii) corresponding to the prime p) there exists a unique polynomial $U(x) \equiv A_{\lambda-N}(x)$, modulo p^N, satisfying (2.2) of minimal degree in x and leading coefficient unity. Let the degree of $A_{\lambda-N}(x)$ be $\alpha_{\lambda-N}$ $(N=1, 2, \cdots, \lambda)$, and let

$$\sigma_N = \alpha_{\lambda-1} + \alpha_{\lambda-2} + \cdots + \alpha_{\lambda-N} \qquad (N = 1, 2, \cdots, \lambda).$$

Then

(iv) *The total number of distinct polynomials modulo* p^N *of degree* $\leqq k-1$ *satisfying* (2.1) *is*

$$p^{Nk-\sigma_N} \textit{ if } N \leqq \lambda \textit{ and } p^{\lambda k-\sigma_\lambda} \textit{ if } N \geqq \lambda.$$

(§6.) In this latter case, the number is therefore independent of N.

3. In the sections which follow, we shall use the German capital $\mathfrak{M}$ for the double modulus $m, f(x)$, writing

$$A(x) \equiv 0 \ (\mathfrak{M}) \text{ for } A(x) \equiv 0 \qquad (\text{modd } m, f(x)).$$

We shall otherwise use the same notation and terminology as in Trans I. In particular, the sequence (U) will be said to be "purely periodic" modulo m if it contains no non-repeating residues when considered modulo m. From Theorem 4.1, Trans I, it suffices to enumerate all the purely periodic sequences (U) with fixed characteristic number s. For if the number of such sequences be denoted by $\psi(s)$, the total number of sequences with characteristic number s may be obtained by multiplying $\psi(s)$ by a factor which is independent of s. (Trans I, part IV.)

By the fundamental theorem on page 606 of Trans I, the enumerative problem for purely periodic sequences is equivalent to the following problem in the theory of congruences to a double modulus:

* Vol. 35 (1933), pp. 254–260. I shall refer to this paper as Trans II.

To determine the total number of distinct polynomials $U(x)$ *modulo* m *of degree* $\leqq k-1$ *such that*

$$U(x)(x^S - 1) \equiv 0 \ (\mathfrak{M}), \tag{3.1}$$

$$U(x)(x^R - 1) \not\equiv 0 \ (\mathfrak{M}) \qquad (1 \leqq R < S). \tag{3.2}$$

We can omit the restriction (3.2). For assume that (3.1) holds, and also that

$$U(x)(x^R - 1) \equiv 0 \ (\mathfrak{M}). \tag{3.3}$$

Then it is easily seen that for any integers L and M,

$$U(x)(x^{LS+MR} - 1) \equiv 0 \ (\mathfrak{M}).$$

Choose L and M so that $LS+MR=D$, the greatest common divisor of S and R. Then

$$U(x)(x^D - 1) \equiv 0 \ (\mathfrak{M}).$$

That is, if (3.3) holds, it must hold for some integer$R = D$ which is a divisor of S. We may therefore replace condition (3.2) by

$$U(x)(x^R - 1) \not\equiv 0 \ (\mathfrak{M}), \ R \text{ any proper divisor of } S. \tag{3.21}$$

Furthermore, if (3.3) holds, there is a smallest value of R for which it holds dividing all other such R.

If $\phi(s)$ is the total number of polynomials $U(x)$ satisfying (3.1) and $\psi(s)$ the total number of polynomials satisfying both (3.1) and (3.21), it is clear then that

$$\phi(s) = \sum_{R \mid S} \psi(R).$$

Therefore by Dedekind's inversion formula,

$$\psi(S) = \sum_{D \mid S} \mu(D)\phi(S/D).$$

The summation here extends over all divisors D of S and $\mu(D)$ denotes Möbius' function. It suffices therefore to determine $\phi(s)$.

4. For the moment, write $u(s;m;f(x))$ for the function $\phi(s)$ defined above. Then first of all, it is readily shown as in Trans I, part III, that if $m=ab$, $(a, b)=1$, then

$$u(s; m; f(x)) = u(s; a; f(x)) \cdot u(s; b; f(x)).$$

That is, $u(s;m;f(x))$ is a multiplicative function of m. *We can assume therefore that*

$$m = p^N, \ p \ \textit{a prime}, \ N \geqq 1. \tag{4.1}$$

Secondly, it is readily shown that if

$$f(x) \equiv f_1(x)\cdot f_2(x) \pmod{m}, \text{ Res } \{f_1(x), f_2(x)\} \text{ prime to } m,$$

then

$$u(s; m; f(x)) = u(s; m; f_1(x))u(s; m; f_2(x)).$$

Since $m = p^N$, we have by Schönemann's second therem* a decomposition of $f(x)$ modulo p^N of the form

$$f(x) \equiv f_1(x)f_2(x) \cdots f_r(x) \pmod{p^N}$$

where $f_i(x)$ is primary and congruent to $\{\phi_i(x)\}^{a_i}$, modulo p, for $i=1$, $\cdot$, $\cdots$, r, while the polynomials $\phi_1(x), \phi_2(x), \cdots, \phi_r(x)$ are distinct and irreducible modulo p.

Since

$$\text{Res } \{f_i(x), f_j(x)\} \not\equiv 0 \pmod{p} \quad (i, j = 1, \cdots, r; i \neq j),$$

we can assume that

$$f(x) = B(x) \equiv \{\phi(x)\}^a \pmod{p},$$

$\phi(x)$ irreducible modulo p.

5. Let

$$B(x) = x^m + b_1x^{m-1} + \cdots + b_m.$$

Then we have reduced our problem to determining the total number of polynomials $U(x)$ of degree $\leqq m-1$, distinct modulo p^N, such that

$$U(x)A(x) \equiv 0 \pmod{p^N, B(x)}, \tag{5.1}$$

where p is a prime, while

$$A(x) \equiv x^S - 1 \pmod{p^N, B(x)}. \tag{5.2}$$

In Trans I, pp. 622–623, I have shown how to determine a polynomial $A(x)$ satisfying (5.2) of degree less than $B(x)$ under the assumption that we know that solution of the difference equation associated with $B(x)$ with the m initial values $0, 0, \cdots, 0, 1$. But here we shall not make any assumption about the degree of $A(x)$. Indeed, we shall show that the number of such polynomials $U(x)$ can be theoretically determined without restricting the form of the polynomial $A(x)$ in any way.

For let

$$A(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n.$$

---

* Fricke, *Algebra*, vol. 2, Braunschweig, 1928, chapter 2.

The congruence (5.1) may be written in the equivalent form

$$A(x)U(x) + B(x)V(x) \equiv 0 \pmod{p^N}, \tag{5.3}$$

where $U(x) = u_0x^{m-1} + \cdots + u_{m-1}$, $V(x) = v_0x^{n-1} + \cdots + v_{n-1}$ are to be determined.

Let $\mathcal{E} = (e_{ij})$ denote the transpose of the matrix corresponding to the Sylvester eliminant of $A(x)$ and $B(x)$. Then if we let

$$z_{i+1} = u_i \quad (i = 0, 1, \cdots, m-1), \quad z_{i+1} = v_{i-m} \; (i = m, m+1, \cdots, m+n-1),$$

(5.3) is equivalent to the set of $n+m$ congruences

$$\sum_{j=1}^{m+n} e_{ij}z_j \equiv 0 \pmod{p^N} \quad (i = 1, 2, \cdots, m+n). \tag{5.4}$$

It is clear then *that the number of distinct polynomials* $U(x)$ *satisfying the conditions of* (5.1) *equals the number of distinct solutions* $z_1, z_2, \cdots, z_{m+n}$ *modulo* p^N *of the system* (5.4).

This number was determined by H. J. S. Smith in a classical memoir.* Namely, let $p^{\lambda_1}, p^{\lambda_2}, \cdots, p^{\lambda_k}$ be the successive elementary divisors of the matrix $\mathcal{E}$ corresponding to the prime p. Then if r is so chosen that $\lambda_{r-1} > N \geqq \lambda_r$, *the number of distinct incongruent solutions of* (5.4) *is* $p^{rN+\lambda_r+\lambda_{r+1}+\cdots+\lambda_k}$.

6. We can express the number of solutions of the congruence (5.1) in quite a different manner by using some of the results obtained in my paper Trans II.

Let us assume that $N \geqq \lambda$, where p^λ now denotes the first elementary divisor of the matrix $\mathcal{E}$ defined in §5 corresponding to p. Then (Trans II, p. 255) $U(x)$ must be of the form

$$U(x) = p^{N-\lambda}(Q_0(x)A_0(x) + pQ_1(x)A_1(x) + \cdots + p^{\lambda-1}Q_{\lambda-1}(x)A_{\lambda-1}(x)), \tag{6}$$

where $A_r(x)$ is the unique polynomial of minimal degree and leading coefficient unity such that

$$A_r(x)A(x) \equiv 0 \qquad (\text{modd } p^{\lambda-r}, B(x)).$$

Let the degree of this polynomial be denoted by α_r.

The procedure by which the polynomials $Q_0(x), Q_1(x), \cdots$ are determined is then as follows:

Let $U(x) = p^{N-\lambda}V_0(x)$. Then $A(x)V_0(x) \equiv 0$ (modd p^λ, $B(x)$) and, as proved in Trans II, $V_0(x) = Q_0(x)A_0(x) + V_1(x)$ where $V_1(x)$ is of lesser degree than

* *On systems of linear indeterminate equations and congruences*, Collected Papers, vol. 1, Oxford, 1894, p. 399.

$A_0(x)$. Then $V_0(x)$ is of degree $\leqq m-1$, $A_0(x)$ is of degree α_0, and $V_1(x)$ of degree $\leqq \alpha_0 - 1$. Hence $Q_0(x)$ is of degree $\leqq m-\alpha_0-1$, so that we can write

$$Q_0(x) = q_1 x^{m-\alpha_0-1} + q_2 x^{m-\alpha_0-2} + \cdots + q_{m-\alpha_0}$$

where $0 \leqq q_j < p^\lambda$ $(j=1, 2, \cdots, m-\alpha_0)$.

Therefore, there are $p^{\lambda(m-\alpha_0)}$ possible polynomials $Q_0(x)$ for a given $U(x)$.

Next, we have

$$A(x)V_1(x) \equiv 0 \qquad (\text{modd } p^{\lambda-1}, B(x)),$$

$V_1(x) = Q_1(x)A_1(x) + pV_2(x)$ where $V_2(x)$ is of lesser degree than $A_1(x)$. Then $V_1(x)$ is of degree $\alpha_0 - 1$, $A_1(x)$ of degree α_1, and $V_2(x)$ of degree $\leqq \alpha_1 - 1$. Therefore $Q_1(x)$ is of degree $\leqq \alpha_0 - \alpha_1 - 1$, and reasoning as before, we see that there are $p^{(\lambda-1)(\alpha_0-\alpha_1)}$ possible polynomials $Q_1(x)$ for a given $U(x)$.

Continuing in this manner, we see that there are $p^{(\lambda-r)(\alpha_{r-1}-\alpha_r)}$ possible polynomials $Q_r(x)$ $(0 \leqq r \leqq \lambda - 1;\ \alpha_{-1} = m)$. Therefore *there are in all*

$$p^{\lambda(m-\alpha_0)} \cdot p^{(\lambda-1)(\alpha_0-\alpha_1)} \cdot p^{(\lambda-2)(\alpha_1-\alpha_2)} \cdot \ \cdots\ \cdot p^{\alpha_{\lambda-2}-\alpha_{\lambda-1}} = p^{\lambda m-(\alpha_0+\alpha_1+\cdots+\alpha_{\lambda-1})}$$

polynomials $U(x)$ satisfying the congruence (5.1); for it easily is seen that each choice of $Q_0(x), \cdots, Q_{\lambda-1}(x)$ in formula (6.1) leads to a distinct polynomial $U(x)$.

If we assume that $N \leqq \lambda$, we have

$$U(x)A(x) \equiv 0 \qquad (\text{modd } p^N, B(x)),$$

$$U(x) = Q_{\lambda-N}(x)A_{\lambda-N}(x) + pV_{\lambda-N+1}(x),$$

$$V_{\lambda-N+1}(x) = Q_{\lambda-N+1}(x)A_{\lambda-N+1}(x) + pV_{\lambda-N+2}(x),$$

and so on.

On determining the degrees of the polynomials $Q_{\lambda-N}(x), Q_{\lambda-N+1}(x), \cdots$, we find that in this case *there are $p^{Nm-(\alpha_{\lambda-N}+\alpha_{\lambda-N+1}+\cdots+\alpha_{\lambda-1})}$ possible polynomials $U(x)$.*

On writing σ_N for $\alpha_{\lambda-1}+\alpha_{\lambda-2}+\cdots+\alpha_{\lambda-N}$ and k for m, we obtain the final result stated in the second section of this paper.

INSTITUTE FOR ADVANCED STUDY,
PRINCETON, N. J.

INTEGRATION IN GENERAL ANALYSIS*

BY

NELSON DUNFORD

L. M. Graves† and others have defined and developed the theory of the Riemann integral for functions whose values are in a complete linear vector space. T. H. Hildebrandt‡ and S. Bochner§ have defined the Lebesgue integral for the same type of function. The present paper, which approaches the theory of the integral in a manner analogous to the Cantor definition of a real number, is concerned chiefly with the convergence of a sequence of integrals and is not as extensive in scope as that of Bochner which contains certain results pertaining to multiple integrals, Fourier series, and the class L_p. In what follows no use is made of the theory of integration for numerically valued functions other than a knowledge of the properties of an additive class of point sets and of a completely additive function on such a class. In fact the method when applied to such functions seems in many ways more direct than the classical one. The proofs in the section on types of convergence are omitted since they may be carried through exactly as in the case of real-valued functions. In the last section it is shown how the theory holds, with slight modifications, for a function having an arbitrary metric space as its domain and a Banach space for its range.

1. **Basis.** A class of point sets is said to be *additive* if for every pair of sets E, D and every sequence $\{E_n\}$ of disjoint sets in the class the sets $E-D$, $\sum E_n$ are also in the class. A function $\alpha(E)$ on an additive class of sets A is said to be *completely additive* if for every sequence $\{E_n\}$ of disjoint sets in A, $\alpha(\sum E_n)=\sum\alpha(E_n)$. In what follows, A will be used to denote an additive class of point sets which contains all Borel measurable sets belonging to a fundamental bounded and closed interval J of a euclidean space of n dimensions. The notation $\alpha(E)$ will always be used for a completely additive function on A to the real number system and $\beta(E)$ will stand for the total variation of α on E. Radon‖ has constructed such systems (A, α, β) corresponding to a

* Presented to the Society, April 20, 1935; received by the editors October 31, 1933, and in revised form, October 30, 1934.

† Graves, *Riemann integration and Taylor's theorem in general analysis*, these Transactions, vol. 29 (1927), pp. 163–177.

‡ Hildebrandt, *Lebesgue integration in general analysis*, this Bulletin, vol. 33 (1927), p. 646.

§ Bochner, *Integration von Funktionen deren Werte die Elemente eines Vektorraumes sind*, Fundamenta Mathematicae, vol. 20 (1933), pp. 262–276.

‖ Radon, *Theorie und Anwendungen der absolut additiven Mengenfunktionen*, Sitzungsberichte der Akademie der Wissenschaften, Wien, Mathematisch-Naturwissenschaftliche Klasse, IIa, vol. 122, pp. 1295–1438.

given function of bounded variation. The properties of α and A which are used below are consequences of their additive nature and not of any particular method of construction, and hence, since we postulate the existence of α, Radon's procedure is not needed in what follows.

2. **Types of convergence.** The functions $f(P)$ to be considered throughout are functions on a subset of J to a complete linear vector space X. The sequence $\{f_n(P)\}$ is said to approach the function $f(P)$ (converge) *almost uniformly with respect to* α *on* E in case for every $\epsilon>0$ there exist sets E_ϵ, E_ϵ' such that E_ϵ' is in A, $E_\epsilon' \supset E-E_\epsilon$, $\beta(E_\epsilon')<\epsilon$ and $f_n(P)\to f(P)$ ($\{f_n(P)\}$ converges) uniformly on E_ϵ. The sequence $\{f_n(P)\}$ is said to approach $f(P)$ *approximately with respect to* α *on* E if for every integer n and every $\epsilon>0$ there exists a set $E'(n,\epsilon)$ in A containing the set $E(n,\epsilon)=E[\|f_n(P)-f(P)\|>\epsilon]$ and such that $\lim_n \beta[E'(n,\epsilon)]=0$. The *convergence of a sequence approximately with respect to* α *on* E is defined similarly by replacing the sets $E'(n,\epsilon)$, $E(n,\epsilon)$ by the sets $E'(m,n,\epsilon)$ and $E(m,n,\epsilon)=E[\|f_m(P)-f_n(P)\|>\epsilon]$ respectively. In what follows, the notations $E(n,\epsilon)$, $E(m,n,\epsilon)$ are used as above. The notation O_α will sometimes be used for a set contained in one on which $\beta=0$. It is assumed that all such sets are in A.

LEMMA 1. *Of the three types of convergence (applied either to the approach to a function or to the convergence of a sequence)*

(1) *almost uniformly with respect to* α *on* E,

(2) *almost everywhere with respect to* α *on* E,

(3) *approximately with respect to* α *on* E,

(1) *implies* (2) *and* (3), *and if* $E(n,\epsilon)$ *and* $E(m,n,\epsilon)$ *are in* A, *then* (2) *implies* (1) *and* (3).

LEMMA 2. *If* E *is in* A, $f_n(P)\to f(P)$ *on* $E-O_\alpha$, *and the set* $E[\|f_n(P)\|>\epsilon]$ *is in* A *for every* n *and* $\epsilon>0$, *then the set* $E[\|f(P)\|>\epsilon]$ *is in* A.

LEMMA 3. *If* $f_n(P)\to f(P)$ *and* $f_n(P)\to f'(P)$ *approximately with respect to* α *on* E *then* $f=f'$ *on* $E-O_\alpha$ *and* $\{f_n(P)\}$ *converges approximately with respect to* α *on* E.

LEMMA 4. *If the sequence* $\{f_n(P)\}$ *converges approximately with respect to* α *on* E, E *being a set of* A, *then there exists a function* $f(P)$ *on* E *to* X *and a subsequence* $\{f_{n_i}(P)\}$ *such that* $f_{n_i}(P)\to f(P)$ *almost uniformly with respect to* α *on* E. *Furthermore if* $\{f_{m_i}(P)\}$ *is any subsequence and* $f^*(P)$ *any function on* E *to* X *such that* $f_{m_i}(P)\to f^*(P)$ *approximately with respect to* α *on* E *then* $f=f^*$ *on* $E-O_\alpha$. *In case the functions* f_m *are uniformly continuous on* E *there exists a set* F *belonging to* A *and contained in* E *such that* F *is closed in* E *and* $\beta(E-F)$ *is arbitrarily small, and* f *as on* F *is uniformly continuous on* F.

Let $U=(u)$, $U'=(u')$, be arbitrary sets of elements and R, R' be transitive order relations on UU, $U'U'$ respectively having the composition property as defined by Moore and Smith.† Then with limits defined in terms of these order relations we have

LEMMA 5. *If $x(u, u')$ is on UU' to X and $\lim_u x(u, u')$ exists for each u' and $\lim_{u'} x(u, u')$ exists uniformly with respect to u then the following limits exist and are equal*: $\lim_{u,u'} x(u, u')$, $\lim_u \lim_{u'} x(u, u')$, $\lim_{u'} \lim_u x(u, u')$.

3. **The integral.** Let $S_0(E)$ denote the class composed of all functions uniformly continuous on E. Let $\pi_E=(E_1, \cdots, E_k)$ represent a partition of the set E which is supposed to be in A, and (P, P') represent the euclidean distance from P to P'. The norm, $n(\pi_E)$, of the partition π_E is defined as the least upper bound of (P_i, P_i') for P_i and P_i' in E_i and for $i=1, \cdots, k$. Then the distance function

$$\|f\| = \int_E \|f(P)\| d\beta = \lim_{n(\pi)=0} \sum_{\pi_E} \|f(P_k)\| \beta(E_k)$$

(P_k being any point in E_k) is surely defined for f in $S_0(E)$. For let

$$S=\sum_{\pi} \|(fP_i)\| \beta(E_i), \quad S'=\sum_{\pi'} \|f(P_k')\| \beta(E_k')$$

be the sums corresponding to the partitions $\pi=(E_i)$, $\pi'=(E_k')$ respectively. Then

$$S = \sum_i \|f(P_i)\| \sum_k \beta(E_iE_k') = \sum_{i,k} \|f(P_i)\| \beta(E_iE_k')$$

and

$$\begin{aligned} |S - S'| &= \Big| \sum_{i,k} (\|f(P_i)\| - \|f(P_k')\|) \beta(E_iE_k') \Big| \\ &\leqq \omega[n(\pi) + n(\pi')] \beta(E) \end{aligned}$$

where

$$\omega[\delta] = \sup_{(P,P') \leqq \delta} \big| \|f(P)\| - \|f(P')\| \big|.$$

Thus the $\lim_{n(\pi)=0} S$ exists for f in $S_0(E)$. The integrals

$$\int_E f(P) d\alpha, \qquad \int_E f(P) d\beta, \qquad \int_E \|f(P)\| d\alpha$$

are defined in a similar manner.

By a *Cauchy sequence* of functions in $S_0(E)$ is meant one for which $\|f_m - f_n\| \to 0$, and two Cauchy sequences $\{f_m\}$ and $\{g_m\}$ are said to be *equivalent* in case $\|f_m - g_m\| \to 0$.

† Moore and Smith, *A general theory of limits*, American Journal of Mathematics, vol. 44 (1922), p. 103.

LEMMA 6. *To each class of equivalent Cauchy sequences of functions in $S_0(E)$ corresponds uniquely except for a set on which $\beta=0$, a function $f(P)$ on E to X such that if $\{f_m\}$ is any sequence of functions in the class defining f then there exists a subsequence $\{f_{m_i}\}$ with $f_{m_i}(P)\to f(P)$ almost uniformly with respect to α on E. Furthermore the limits*

$$\lim_n \int_E f_n(P)d\alpha, \qquad \lim_n \int_E f_n(P)d\beta,$$

$$\lim_n \int_E \|f_n(P)\|d\alpha, \qquad \lim_n \int_E \|f_n(P)\|d\beta$$

all exist and are independent of the particular Cauchy sequence in the class of equivalent Cauchy sequences.

The set $E(m, n, \epsilon)$ is the product of a region and the set E, and is thus in A. Now

$$\begin{aligned} \epsilon\beta[E(m, n, \epsilon)] &\leqq \int_{E(m,n,\epsilon)} \|f_m(P) - f_n(P)\|d\beta \\ &\leqq \int_E \|f_m(P) - f_n(P)\|d\beta \to 0 \end{aligned}$$

so that $\{f_m\}$ converges approximately with respect to α on E. The existence of $f(P)$ follows from Lemma 4, and its uniqueness may be established in a manner similar to that used in the proof of Lemma 4. Since

$$\left\| \int_E f_n(P)d\alpha - \int_E f_m(P)d\alpha \right\| \leqq \int_E \|f_m(P) - f_n(P)\|d\beta \to 0,$$

it follows that the $\lim_n \int_E f_n(P)d\alpha$ exists. It is independent of the sequence $\{f_m\}$ since

$$\left\| \int_E f_n(P)d\alpha - \int_E g_n(P)d\alpha \right\| \leqq \|f_n - g_n\| \to 0$$

if $\{f_n\}$ and $\{g_n\}$ are equivalent. In the same manner the other limits may be shown to exist.

A function $f(P)$ is said to be *summable with respect to α on E* in case it is the correspondent in the sense of Lemma 6 of some class of equivalent Cauchy sequences of functions in $S_0(E)$. The class of such functions will be denoted by $S(E)$. The *integral* $\int_E f(P)d\alpha$ of a function in $S(E)$ is defined as the $\lim_n \int_E f_n(P)d\alpha$ where $\{f_n(P)\}$ is any sequence in the class defining f. The integrals $\int_E \|f(P)\|d\alpha$, $\int_E \|f(P)\|d\beta$, $\int_E f(P)d\beta$ are defined similarly.

Note that $\|\int_E f(P)d\alpha\| \leqq \int_E \|f(P)\|d\beta$.

THEOREM 1. *If f is in $S(E)$ then the set $E[\|f(P)\|>\epsilon]$ is in A for every $\epsilon>0$.*

This is a corollary of Lemmas 1, 2.

THEOREM 2. *Every function $f(P)$ in $S(E)$ is approachable almost uniformly with respect to α on E by a sequence $f_m(P)$ of functions uniformly continuous on E and such that $\int_E\|f_m(P)-f(P)\|d\beta\to 0$.*

The first part of the conclusion is a corollary of Lemma 6, and the second part follows from the fact that for a fixed value of m the sequence $f_m(P)-f_n(P)$ is a Cauchy sequence of functions in $S_0(E)$ belonging to the class defining $f_m(P)-f(P)$. Thus

$$\int_E\|f_m(P) - f(P)\|d\beta = \lim_n \int_E \|f_m(P) - f_n(P)\|d\beta$$

and the conclusion is immediate.

THEOREM 3. *If f is in $S(E)$ and $\epsilon>0$, then there exists a set F belonging to A and contained in E such that F is closed in E, f as on F is uniformly continuous on F and $\beta(F)\geqq\beta(E)-\epsilon$.*

This is a corollary of Theorem 2, Lemma 1, and Lemma 4.

THEOREM 4. *The space $S(E)$ of functions summable with respect to α on E with the norming operation $\|f\|=\int_E\|f(P)\|d\beta$ is a complete linear vector space.*

Let $\{f_m\}$ be a Cauchy sequence of functions in $S(E)$; then by Theorem 2, for every m there exists a function g_m in $S_0(E)$ such that $\|f_m-g_m\|<1/m$. Now

$$\|g_m - g_n\| \leqq \|g_m - f_m\| + \|f_m - f_n\| + \|f_n - g_n\|,$$

so that $\{g_m\}$ is a Cauchy sequence of functions in $S_0(E)$ and thus defines a function $f(P)$ in $S(E)$ such that $\|g_n-f\|\to 0$. Thus

$$\|f_m - f\| \leqq \|f_m - g_m\| + \|g_m - f\| \to 0,$$

and $S(E)$ is complete. The rest of the proof follows immediately from the definition of the terms involved.

A function $h(E)$ on an additive class contained in A to a linear vector space is said to be *absolutely continuous with respect to* α in case $\lim_{\beta(E)=0} h(E)=0$.

THEOREM 5. *If $f(P)$ is in $S(E)$ then the integrals $\int_e f(P)d\alpha$, $\int_e f(P)d\beta$, $\int_e\|f(P)\|d\alpha$, $\int_e\|f(P)\|d\beta$ are all absolutely continuous with respect to α.*

If $f(P)$ is in $S_0(E)$ then $\|f(P)\|$ is bounded on E and so

$$\left\| \int_e f(P)d\alpha \right\| \leqq M\beta(e).$$

Now if f_m is a Cauchy sequence in $S_0(E)$ with $\|f_m - f\| \to 0$, we have for each m

$$\lim_{\beta(e)=0} \int_e f_m(P)d\alpha = 0,$$

and since

$$\left\| \int_e (f_m(P) - f(P))d\alpha \right\| \leqq \|f_m - f\| \to 0,$$

it follows that

$$\lim_m \int_e f_m(P)d\alpha = \int_e f(P)d\alpha \text{ uniformly with respect to } e.$$

Thus by Lemma 5

$$\lim_{\beta(e)=0} \int_e f(P)d\alpha = 0.$$

THEOREM 6. *If $f(P)$ is in $S(E)$ then each of the integrals listed in Theorem 5 is a completely additive function on the class $A(E)$ composed of all sets e in A and such that $e \subset E$.*

If $f(P)$ is in $S_0(E)$ and e_i, $i=1, \cdots, k$, are disjoint sets in $A(E)$, it follows from the definition of the integral on the class $S_0(E)$ that

$$\int_{\sum_{i=1}^k e_i} f(P)d\alpha = \sum_{i=1}^k \int_{e_i} f(P)d\alpha.$$

Thus for any $f(P)$ in $S(E)$ the same equation holds. Now for a sequence $\{e_i\}$ of disjoint sets with $\sum e_i = e$,

$$\left\| \int_e f(P)d\alpha - \sum_{i=1}^m \int_{e_i} f(P)d\alpha \right\| = \left\| \int_{e-\sum_{i=1}^m e_i} f(P)d\alpha \right\| \to 0$$

by Theorem 5.

THEOREM 7. *If f and f_m are in $S(E)$ for every integer m and if $\|f_m - f\| \to 0$, then $f_m \to f$ approximately with respect to α on E, $\|f_m\|$ is bounded, and $\int_e \|f_m(P)\| d\beta$ is absolutely continuous with respect to α uniformly with respect to m.*

Since $f_m - f$ is in $S(E)$ the set $E(m, \epsilon)$ is by Theorem 1 in A. Now

$$\epsilon\beta[E(m, \epsilon)] \leqq \int_{E(m,\epsilon)} \|f_m(P) - f(P)\| d\beta \leqq \|f_m - f\| \to 0$$

so that $f_m \to f$ approximately with respect to α on E. Also since $\|f_m\| \leqq \|f_m - f\| + \|f\|$, $\|f_m\|$ is bounded. The remaining part of the conclusion follows from the inequality

$$\int_e \|f_n(P)\| d\beta \leqq \|f_n - f\| + \int_e \|f(P)\| d\beta$$

and Theorem 6.

THEOREM 8. *If $\{f_m(P)\}$ is a sequence of functions in $S(E)$ and $f_n(P) \to f(P)$ approximately with respect to α on E, and if $\int_e \|f_m(P)\| d\beta$ is absolutely continuous with respect to α uniformly with respect to m, then f is in $S(E)$ and $\int_e \|f_n(P) - f(P)\| d\beta \to 0$ uniformly for e in $A(E)$.*

By Lemma 3 the sequence $\{f_m(P)\}$ converges approximately. By Theorem 1, $E(m, n, \epsilon)$ is in A. Now

$$\int_E \|f_m(P) - f_n(P)\| d\beta = \int_{E-E(m,n,\epsilon)} + \int_{E(m,n,\epsilon)}$$

$$\leqq \epsilon\beta[E - E(m, n, \epsilon)] + \int_{E(m,n,\epsilon)}.$$

Thus $\|f_m - f_n\| \to 0$ and by Theorem 4 there exists a function f' such that $\|f_m - f'\| \to 0$. By Theorem 7 and Lemma 3 it is seen that $f = f'$ on $E - O_\alpha$ and thus f is in $S(E)$. Since $\int_e \|f_m(P) - f(P)\| d\beta \leqq \|f_m - f\|$ the proof is complete.

COROLLARY 1. *If the sequence $\{f_n(P)\}$ of functions in $S(E)$ approach $f(P)$ approximately with respect to α on E, and if there exists a function $g(P)$ in $S(E)$ such that $\|f_n(P)\| \leqq \|g(P)\|$ on $E - O_\alpha$ for every n, then f is in $S(E)$ and $\int_e \|f_m(P) - f(P)\| d\beta \to 0$ uniformly for e in $A(E)$.*

COROLLARY 2. *If f is in $S(E)$ and $\phi(P)$ is a real-valued function summable with respect to α on E and bounded on $E - O_\alpha$, then $\phi(P)f(P)$ is in $S(E)$.*

THEOREM 9. *If f_m and f are in $S(E)$ and $f_m \to f$ approximately with respect to α on E, then the following assertions are equivalent:*

(1) $\lim_m \int_e f_m d\alpha = \int_e f d\alpha$ *on* $A(E)$.

(2) $\lim_m \int_e f_m d\alpha = \int_e f d\alpha$ *uniformly on* $A(E)$.

(3) *The* $\lim_m \int_e f_m d\alpha$ *exists on* $A(E)$.

(4) $\lim_{\beta=0} \overline{\lim}_m \|\int_e f_m d\alpha\| = 0$.

(5) $\lim_{\beta=0} \int_e f_m d\alpha = 0$ *uniformly*.

The proof will be made by demonstrating the following implications: (4)→(1)→(3)→(5)→(2)→(4), where each arrow is directed from hypothesis to conclusion. To see that (4)→(1) first note that (4) is equivalent to the

statement that for every $\epsilon>0$ there exists a $\delta>0$ such that for each e in $A(E)$ with $\beta(e)<\delta$ there is an n_e such that $\|\int_e f_n d\alpha\|<\epsilon$ for all $n\geqq n_e$. Now let it be supposed that $f_n\rightarrow f$ almost uniformly with respect to α on E. To show that $\int_e f_m d\alpha\rightarrow\int_e f d\alpha$, note that

$$\left\|\int_E (f_m - f)d\alpha\right\| \leqq \left\|\int_{E-e} (f_m - f)d\alpha\right\| + \left\|\int_e (f_m - f)d\alpha\right\|.$$

Fix e such that $\beta(E-e)<\delta$ and $f_m\rightarrow f$ uniformly on e. For this e there is an m_1 such that

$$\left\|\int_{E-e} (f_m(P) - f(P))d\alpha\right\| < \epsilon, \quad \left\|\int_e (f_m(P) - f(P))d\alpha\right\| < \epsilon, \quad m \geqq m_1.$$

Thus $\int_E f_m d\alpha\rightarrow\int_E f d\alpha$. The same proof holds for an arbitrary e in $A(E)$. The same conclusion follows if $f_n\rightarrow f$ approximately since by Lemma 4 and the above proof every subsequence of $\{\int_e f_m d\alpha\}$ contains a subsequence approaching $\int_e f d\alpha$. That (1)→(3) is obvious. For the proof of the implication (3)→(5) we refer to a paper by Saks.* Now (5)→(2), for

$$\left\|\int_e (f_m - f)d\alpha\right\| \leqq \left\|\int_{e(m,\epsilon)} (f_m - f)d\alpha\right\| + \left\|\int_{e-e(m,\epsilon)} (f_m - f)d\alpha\right\|.$$

Now

$$\left\|\int_{e-e(m,\epsilon)} (f_m - f)d\alpha\right\| \leqq \epsilon\beta(E),$$

and there exists an m_0 and a $\delta>0$ such that $\beta[E(m,\epsilon)]<\delta$ for $m\geqq m_0$ and

$$\left\|\int_e (f_m(P) - f(P))d\alpha\right\| < \epsilon$$

for all m and e with $\beta(e)<\delta$. Thus since $e(m,\epsilon)\subset E(m,\epsilon)$ it follows that

$$\left\|\int_e (f_m(P) - f(P))d\alpha\right\| \leqq \epsilon(1 + \beta(E)) \text{ for } m \geqq m_0.$$

Finally in view of Theorem 6 and the fact that (2) implies that

$$\lim_m \left\|\int_e f_m d\alpha\right\| = \left\|\int_e f d\alpha\right\|,$$

the implication (2)→(4) is obvious.

---

* *Addition to the note on some functionals*, these Transactions, vol. 35 (1933), p. 967.

THEOREM 10. *If E is in A; α_m, α are completely additive functions on $A(E)$; $\alpha_m(E_i^k) \to \alpha(E_i^k)$ for each set E_i^k found in a sequence of partitions $\pi_i(E) = (E_i^k)$ with $n(\pi_i) \to 0$; and if the sequence $\{\beta_m(E)\}$ is bounded; then*

$$\lim_m \int_E f d\alpha_m = \int_E f d\alpha$$

for any function $f(P)$ in $S_0(E)$.

From the proof given at the beginning of §3 and the boundedness of the sequence $\{\beta_m(E)\}$ it follows that

$$\lim_{i=\infty} \sum_k f(P_i^k)\alpha_m(E_i^k) = \int_E f d\alpha_m$$

uniformly with respect to m. Also for each i

$$\lim_m \sum_k f(P_i^k)\alpha_m(E_i^k) = \sum_k f(P_i^k)\alpha(E_i^k).$$

The conclusion follows from Lemma 5.

THEOREM 11. *If E is in A; α_m, α are completely additive on $A(E)$; $f(P)$ summable with respect to α_m and α on E; $\alpha_n(e) \to \alpha(e)$ on $A(E)$; and if the sequence $\{\beta_m(E)\}$ is bounded; then*

$$\int_e f d\alpha_m \to \int_e f d\alpha \text{ on } A(E)$$

provided

$$\lim_{\beta(e)=0} \overline{\lim_{m=\infty}} \left\| \int_e f d\alpha_m \right\| = 0.$$

We have

$$\left\| \int_E f d\alpha_m - \int_E f d\alpha \right\| \leqq \left\| \int_{E-e} f d\alpha_m \right\| + \left\| \int_e f d\alpha_m - \int_e f d\alpha \right\| + \left\| \int_{E-e} f d\alpha \right\|.$$

Now for $\epsilon > 0$ there is a $\delta > 0$ such that for every $E - e$ with $\beta(E-e) < \delta$ there is an integer m_0 such that $m \geqq m_0$ implies

$$\left\| \int_{E-e} f d\alpha_m \right\| < \epsilon, \qquad \left\| \int_{E-e} f d\alpha \right\| < \epsilon.$$

Fix e with $\beta(E-e) < \delta$ so that f as on e is uniformly continuous on e. Then for m sufficiently large (Theorem 10)

$$\left\| \int_e f d\alpha_m - \int_e f d\alpha \right\| < \epsilon$$

and thus for m sufficiently large

$$\left\| \int_E f d\alpha_m - \int_E f d\alpha \right\| < 3\epsilon.$$

The conclusion follows for an arbitrary e in $A(E)$ since all the hypotheses are satisfied by e when they are by E.

COROLLARY. *Suppose E in A; α_m, α, γ completely additive on $A(E)$; $\alpha_m(e) \to \alpha(e)$ on $A(E)$; $\beta_m(e) \leqq \gamma(e)$ on $A(E)$ for every m; and f is summable with respect to γ on E; then f is summable with respect to α_m, α on E and $\int_e f d\alpha_m \to \int_e f d\alpha$ on $A(E)$.*

4. **The generalization to the case where J is a metric space.** The above theory of the integral holds if the domain of the function $f(P)$ is a general metric space rather than a euclidean space of n dimensions. The few alterations and additions in the arguments that are necessary will be enumerated here.

J will now be interpreted as an arbitrary metric space not necessarily of bounded diameter. The class $S_0(E)$ will be the class of all functions $f(P)$ *uniformly continuous and bounded on E*. By a partition of the set E in J will be meant a set (E_η) of disjoint sets (possibly non-denumerable in number) such that $E = \sum E_\eta$ and which is found in the following manner. Let ϵ be an arbitrary positive number and $(P)_\epsilon$ denote all points P' in E for which $(P, P') < \epsilon$. Take any point P_1 in E; then $E_1 = (P_1)_\epsilon$. In general P_η is any point in $E - \sum_{\zeta<\eta} E_\zeta$ and $E_\eta = (P_\eta)_\epsilon - \sum_{\zeta<\eta} E_\zeta$. The sets E_η form a partition $\pi(E)$ of E with $n(\pi(E)) \leqq 2\epsilon$. It should be mentioned perhaps that such partitions will only be allowed in the definition to be given of $\int_E f(P) d\alpha$ and not in the definition of $\beta(E)$, the latter remaining unaltered. The partition just defined is devised to avoid assuming the separability of J as well as to eliminate the possibility of using a partition of certain connected sets each set of which consists of a single point.

Since β is completely additive, $\beta(E_\eta) = 0$ for all except at most a denumerable number of the sets E_η in any partition $\pi(E) = (E_\eta)$. This follows from the fact that for an arbitrary integer m there can be at most a finite number of the sets E_η for which $\beta(E_\eta) > 1/m$. If those sets E_η of the partition $\pi(E)$ for which $\beta(E_\eta) \neq 0$ are arranged into a sequence $\{E_i\}$ in *any order*, it is possible to associate with the partition $\pi(E)$ a sum

$$\sum_{i=1}^{\infty} f(P_i)\alpha(E_i)$$

where P_i now stands for any point in E_i. If $f(P)$ is in $S_0(E)$, the above sum exists and is independent of the particular arrangement of the terms in the sequence $\{E_i\}$. First the sum exists for any particular arrangement, since for $m' > m$

$$\left\| \sum_{i=m}^{m'} f(P_i)\alpha(E_i) \right\| \leqq \sup_{P \epsilon E} \|f(P)\| \sum_{i=m}^{m'} \beta(E_i).$$

Now let

$$x_1 = \sum_{i=1}^{\infty} f(P_i)\alpha(E_i), \qquad x_2 = \sum_{i=1}^{\infty} f(P_i')\alpha(E_i')$$

be the sums for two different arrangements of the sequence $\{E_i\}$. For every $\epsilon > 0$ there is an integer m_1 such that, for $m \geqq m_1$,

$$\left\| x_1 - \sum_{i=1}^{m} f(P_i)\alpha(E_i) \right\| < \epsilon,$$

$$\left\| x_2 - \sum_{i=1}^{m} f(P_i')\alpha(E_i') \right\| < \epsilon,$$

$$\sum_{i=m_1}^{\infty} \beta(E_i) < \epsilon / \sup_{P \epsilon E} \|f(P)\|.$$

Now suppose m_2 the largest value of i for which the set E_i' is one of the sets $E_1, \cdots, E_{m_1}$; and E_k'', $k = 1, \cdots, m_2 - m_1$, are those E_i' $(1 \leqq i \leqq m_2)$ which are not found among the sets $E_1, \cdots, E_{m_1}$. Then $m_2 \geqq m_1$ and

$$\left\| \sum_{i=1}^{m_2} f(P_i')\alpha(E_i') - \sum_{i=1}^{m_1} f(P_i)\alpha(E_i) \right\| = \left\| \sum_{k=1}^{m_2 - m_1} f(P_k'')\alpha(E_k'') \right\|$$

$$\leqq \sup_{P \epsilon E} \|f(P)\| \cdot \sum_{i=m_1}^{\infty} \beta(E_i) \leqq \epsilon.$$

Thus $\|x_1 - x_2\| \leqq 3\epsilon$, $x_2 = x_1$.

The integral $\int_E f(P)d\alpha$ is now defined as

$$\lim_{n(\pi)=0} \sum_{\pi(E)} f(P_i)\alpha(E_i).$$

The proof given in the text for the existence of this limit for f in $S_0(E)$ holds verbatim with the additional point involved in the justification of the equality

$$\sum_i f(P_i) \sum_k \alpha(E_i E_k') = \sum_{i,k} f(P_i)\alpha(E_i E_k').$$

This equality is established immediately with the use of Lemma 5.

In the proof of Theorem 6 another argument must be added. It is necessary to show that if $f(P)$ is in $S_0(E)$ and e', e'' are disjoint subsets of E, then

$$\sum_i f(P_i)\alpha(e_i) = \sum_i f(P_i')\alpha(e_i') + \sum_i f(P_i'')\alpha(e_i''),$$

where e_i' (e_i'') are those sets of a partition of $e'(e'')$ on which $\beta \neq 0$ and the partition of the set $e = e' + e''$ is formed by a combination of the two partitions. Let

$$x = \sum_{i=1}^{\infty} f(P_i)\alpha(e_i), \qquad x' = \sum_{i=1}^{\infty} f(P_i')\alpha(e_i'), \qquad x'' = \sum_{i=1}^{\infty} f(P_i'')\alpha(e_i'');$$

then for $\epsilon > 0$ there is an m_1 such that, for $m \geqq m_1$,

$$\left\| \sum_{i=1}^{m} f(P_i)\alpha(e_i) - x \right\| < \epsilon,$$

$$\left\| \sum_{i=1}^{m} f(P_i')\alpha(e_i') - x' \right\| < \epsilon,$$

$$\left\| \sum_{i=1}^{m} f(P_i'')\alpha(e_i'') - x'' \right\| < \epsilon,$$

$$\sum_{i=m_1}^{\infty} \beta(e_i' + e_i'') < \epsilon / \sup_{P \epsilon E} \|f(P)\|.$$

Let m_2 be the largest value of i for which the set e_i is one of the sets e_i', e_i'', $i = 1, \cdots, m_1$, and let e_i''', $i = 1, \cdots, m_2 - 2m_1$, be those of the sets e_i, $i = 1, \cdots, m_2$, which are not found among the sets e_i', e_i'', $i = 1, \cdots, m_1$; then

$$\left\| \sum_{i=1}^{m_1} \left[f(P_i')\alpha(e_i') + f(P_i'')\alpha(e_i'') \right] - \sum_{i=1}^{m_2} f(P_i)\alpha(e_i) \right\|$$
$$= \left\| \sum_{i=1}^{m_2 - 2m_1} f(P_i''')\alpha(e_i''') \right\| \leqq \sup_{P \epsilon E} \|f(P)\| \sum_{i=m_1}^{\infty} \beta(e_i' + e_i'') < \epsilon,$$

and so $\|x' + x'' - x\| < 4\epsilon$, $x = x' + x''$.

We see no way of proving Theorem 10 unless the hypothesis is strengthened to the extent that $\alpha_m \rightarrow \alpha$ on $A(E)$. It is then possible to show that

$$\lim_m \sum_\pi f(P_i)\alpha_m(E_i) = \sum_\pi f(P_i)\alpha(E_i)$$

for any partition π of the set E, where now the $\sum_\pi$ is to be taken over all sets E_η of the partition π for which any one of the inequalities $\beta(E_\eta)\neq 0$, $\beta_m(E_\eta)\neq 0$ hold. We have

$$\lim_m \sum_{i=1}^n f(P_i)\alpha_m(E_i) = \sum_{i=1}^n f(P_i)\alpha(E_i)$$

for each n, and

$$\lim_n \sum_{i=1}^n f(P_i)\alpha_m(E_i) = \sum_{i=1}^\infty f(P_i)\alpha_m(E_i)$$

uniformly with respect to m. To see this, note that the functions α_m are absolutely continuous with respect to the completely additive function

$$\gamma(e) = \sum_{n=1}^\infty \frac{\beta_n(e)}{2^n(\beta_n(E)+1)},$$

and thus by a theorem of Saks* they are absolutely continuous uniformly with respect to m. Thus for $\epsilon>0$ there is a $\delta>0$ such that

$$|\alpha_m(e)| < (\epsilon/2)\sup_{P\epsilon E}\|f(P)\|$$

for every m provided $\gamma(e)<\delta$. Now there is an n_1 such that

$$\sum_{i=n_1}^\infty \gamma(E_i) = \gamma\left(\sum_{i=n_1}^\infty E_i\right) < \delta.$$

If s_+ denotes the set of all integers $i\geqq n_1$ for which $\alpha_m(E_i)\geqq 0$ and s_- those $i\geqq n_1$ for which $\alpha_m(E_i)<0$,

$$\sum_{i=n_1}^\infty |\alpha_m(E_i)| = \alpha_m\left(\sum_{s+} E_i\right) - \alpha_m\left(\sum_{s-} E_i\right) \leqq \epsilon/\sup_{P\epsilon E}\|f(P)\|,$$

and for $n\geqq n_1$,

$$\left\|\sum_{i=n}^\infty f(P_i)\alpha_m(E_i)\right\| \leqq \sup_{P\epsilon E}\|f(P)\|\sum_{i=n_1}^\infty |\alpha_m(E_i)| \leqq \epsilon.$$

Thus, by Lemma 5,

$$\lim_m \sum_\pi f(P_i)\alpha_m(E_i) = \sum_\pi f(P_i)\alpha(E_i).$$

* Saks, loc. cit.

BROWN UNIVERSITY,
PROVIDENCE, R. I.

ON CYCLIC FIELDS*

BY

A. ADRIAN ALBERT

1. **Introduction.** The most interesting algebraic extensions of an arbitrary field F are the cyclic extension fields Z of degree n over F. I have recently given constructions of such fields for the case $n=p$,† a prime, when the characteristic of F is not p, and for the case $n=p^e$‡ when the characteristic of F is p. Moreover it is well known that when F contains all the nth roots of unity then $Z=F(x)$, $x^n=\alpha$ in F.

The last result above does not provide a construction of all cyclic fields Z over F since in general F does not contain these nth roots. Moreover if we adjoin these roots to F and so extend F to a field K the composite (Z, K) over K may not have degree §n. Finally even if (Z, K) over K does have degree n then it is necessary to give conditions that a given field $K(x)$, $x^n=\alpha$ in F, shall have the form (Z, K) with Z cyclic over F. This has not been done and is certainly not as simple as the considerations I shall make here.

It is well known that if $n=p_1^{e_1}\cdots p_t^{e_t}$ with p_i distinct primes, then Z is the direct product $Z=Z_1\times\cdots\times Z_t$ where Z_i is cyclic of degree p over F. Hence it suffices to consider the case $n=p^e$, p a prime. I have already done so‡ for the case where F has characteristic p. In the present paper I shall make analogous considerations for the case where F has characteristic not p by first studying the case where F contains a primitive pth root of unity ζ and later giving complete conditions for the case where F does not contain ζ.

2. **Algebraic units of Z.** Let Z be cyclic of degree n over a field F and S be a generating automorphism of the automorphism group of Z. Then we define the relative norm

$$N_{Z/F}(a) = aa^S\cdots a^{S^{n-1}}, \tag{1}$$

a quantity of F for every a of Z. We shall now give a new proof of a theorem of Hilbert.||

* Presented to the Society, September 7, 1934; received by the editors July 30, 1934.

† See my paper in these Transactions, 1934, *On normal Kummer fields over a non-modular field.* The results and proofs hold if F is any field of characteristic not p.

‡ Bulletin of the American Mathematical Society, vol. 40 (1934), pp. 625–631.

§ For let Z be the field of the 2^{n+1} roots of unity so that Z has degree 2^n over R, the rational field. Then K is actually a sub-field of degree 2^{n-1} of Z and Z has degree 2 over K.

|| Cf. Hilbert's *Abhandlungen* I, p. 149. Hilbert's proof uses the assumption that F is infinite and is very different from the rather interesting proof given here. The proof here also goes more deeply into the true reason for the theorem.

THEOREM 1. *A quantity a of Z has the property*

$$N_{Z/F}(a) = 1 \tag{2}$$

if and only if there exists a quantity $b \neq 0$ of Z such that

$$a = b^S/b. \tag{3}$$

For obviously if a has the form (3) then $N_{Z/F}(a) = N_{Z/F}(b)N_{Z/F}(b^{-1}) = 1$. Conversely let $N_{Z/F}(a) = 1$.

Consider the cyclic algebra M whose quantities are all $\sum_{i=0}^{n-1} z_i y^i$ with z_i in Z and $1, y, \cdots, y^{n-1}$ left linearly independent in Z. Let

$$y^i z = z^{S^i} y^i, \quad y^n = 1 \qquad (z \text{ in } Z), \tag{4}$$

so that M is *equivalent to the algebra of all n-rowed square matrices*. Then Z may be thought of as a field of n-rowed square matrices, y is a matrix whose minimum equation is $y^n - 1 = 0$, its characteristic equation. The matrix $a^{-1}y = y_0$ has the property $y_0^n = N(a^{-1}) = 1$ and has the same minimum equation as y. Since this equation defines the only invariant factor of y which is not unity, the two matrices y and y_0 have the same invariant factors and are similar. Thus $y_0 = AyA^{-1}$ with $A = \sum z_i y^i \neq 0$ and

$$yA = aAy = \sum z_i^S y^{i+1} = a \sum z_i y^{i+1}.$$

Then $az_i = z_i^S \neq 0$ for at least one z_i so that we take $b = z_i \neq 0$.

3. **Cyclic fields of degree p^e over K.** Let K be a field of characteristic not p containing a primitive pth root of unity ζ and let Z be cyclic of degree p^e over K, $e > 1$. Then Z contains a unique cyclic sub-field Y of degree $m = p^{e-1}$ and Z is cyclic of degree p over Y. But then*

$$Z = Y(z), \quad z^p = a \text{ in } Y. \tag{5}$$

Let S be a generating automorphism of Z so that S may also be considered as a generating automorphism of Y. Then $S^m = Q$, $Q^p = I$, the identity automorphism of Z, and Y is the set of all quantities of Z unaltered by the cyclic group $(I, Q, \cdots, Q^{p-1})$.

We compute $(z^Q)^p = a^Q = a$. Then z^Q is a root of $\omega^p = a$ and hence

$$z^Q = \zeta^\mu z \qquad (0 \leqq \mu < p). \tag{6}$$

If $\mu = 0$ then $z^Q = z$ is in Y contrary to our hypothesis that $Z = Y(z) \neq Y$. Hence $\mu > 0$ is prime to p,

$$\mu\mu_0 = 1 + \mu_1 p, \qquad (\mu_0, p) = 1, \tag{7}$$

for integers μ_0, μ_1. Define $S_0 = S^{\mu_0}$, $Q_0 = Q^{\mu_0}$ so that S_0 is a generating auto-

* For every cyclic field of degree p over Y containing ζ is a Kummer field $Y(z)$, $z^p = a$ in Y.

morphism of Z, Q_0 is a generator of the group $(I, Q, \cdots, Q^{p-1})$. Then $z^{Q_0} = \zeta^{\mu\mu_0} z = \zeta z$. Hence by properly choosing S we may assume

$$z^Q = \zeta z, \tag{8}$$

instead of (6).

Now $(z^S)^p = a^S$ so that, by a well known theorem on Kummer fields,* we have $z^S = \beta z^\nu$, β in Y, $1 \leqq \nu < p$. Then

$$z^{S^2} = \beta^S \beta^\nu z^{\nu^2} = \beta_2 z^{\nu^2}, \cdots, z^{S^m} = \beta_\nu z^{\nu^m} = z^Q = \zeta z$$

and hence $z^{\nu^m - 1} = \beta_\nu^{-1}\zeta$ is in the field Y. But then $\nu^m \equiv 1 \pmod p$ and, since $m = p^{e-1}$ so that $\nu^m \equiv \nu \pmod p$ we have $\nu \equiv 1 \pmod p$, $\nu = 1$.

Then

$$z^S = \beta z, \quad \beta \text{ in } Y. \tag{9}$$

Also

$$z^{S^2} = \beta^S \beta z, \cdots, z^{S^m} = z^Q = \beta^{S^{m-1}} \cdots \beta^S \beta z$$

and

$$N_{Y/K}(\beta) = \zeta. \tag{10}$$

The quantity β is in Y and has the property (10) so that $N_{Z/K}(\beta) = N_{Y/K}(\beta^p) = \zeta^p = 1$. By Theorem 1 applied in Y we have

$$\beta^p = \frac{\alpha^S}{\alpha}, \quad \alpha \text{ in } Y. \tag{11}$$

But now $a^S = (z^S)^p = \beta^p a$ so that

$$(\alpha a^{-1})^S = \alpha a^{-1}, \tag{12}$$

and hence $\alpha = \lambda a$ with λ in K.

We may finally prove that in fact $Z = K(z)$. This will obviously be true if $z^p = a$ generates Y. Hence let a be in a proper sub-field of Y. Then a is in the unique sub-field H of degree p^{e-2} of Y and if $m = pr$, $R = S^r$, we have $R^p = Q$, $a^R = a$. Then $a^S = a\beta^p$, $a^R = a(\beta\beta^S \cdots \beta^{S^{r-1}})^p = a$ so that $[N_{H/K}(\beta)]^p = 1$, $N_{H/K}(\beta) = \zeta^\rho$, $N_{Y/K}(\beta) = \zeta^{\rho p} = 1$, a contradiction. We have proved

THEOREM 2. *Let Z be a cyclic field of degree p^e over K, $e > 1$, S be a generating automorphism of Z, and Y its unique sub-field of degree p^{e-1} over K. Then $Z = K(z)$ where $z^p = a$ in Y and Y contains a quantity β such that*

$$N_{Y/K}(\beta) = \zeta, \quad a^S a^{-1} = \beta^p. \tag{13}$$

---

* Cf. Hasse's *Bericht über Klassenkörper*, Jahresbericht der Deutschen Mathematiker Vereinigung, vol. 36 (1927), pp. 233–311; p. 262.

Moreover the generating automorphism S of Z is given by that in Y and

$$z^S = \beta z. \tag{14}$$

We may now prove

THEOREM 3. *A necessary and sufficient condition that a cyclic field Y of degree p^{e-1} over K, $e>1$, shall possess cyclic overfields of degree p^e over K is that Y shall contain a quantity β such that $N_{Z/K}(\beta)=\zeta$. Every such cyclic overfield* is a field $K(z)$, $z^p=a_0$, with generating automorphism* (14), *where $a_0=\lambda a$, a is any root of*

$$a^S a^{-1} = \beta^p, \tag{15}$$

and λ ranges over all quantities of K.

For if Z is cyclic of degree p^e over K then the existence of β is given by Theorem 2. Conversely let $N_{Z/K}(\beta)=\zeta$ for β in Y. By Theorem 1 there exists a quantity a in Y such that (15) is satisfied. If $a=b^p$ for b in K then $a^S a^{-1}=(b^S b^{-1})^p=\beta^p$, $\beta=\zeta^r b^S b^{-1}$, $N_{Y/K}(\beta)=1$, a contradiction. Hence the field $Z=Y(z)$, $z^p=a_0$, has degree p over Y for every solution a_0 of $a^S a^{-1}=\beta^p$. Moreover $a_0=\lambda a$ for any fixed solution a. In our proof of Theorem 2 we showed that in fact $Y=K(a_0)$ so that $Z=K(z)$. Finally Z is evidently a field of Theorem 2 and is cyclic with generating automorphism given by that in Y and by (14).

Suppose now that Z_0 is a new cyclic overfield of Y of degree p^e over K so that Z_0 defines a quantity β_0 with $N_{Y/K}(\beta_0)=\zeta$. Then $N_{Y/K}(\beta_0\beta^{-1})=1$ and

$$\beta_0 = \beta d^S d^{-1}, \tag{16}$$

with d in Y by Theorem 1. Moreover $Z_0=K(z_1)$, $z_1^p=a_1$, where $a_1^S a_1^{-1}=\beta_0^p$. But if $a_{01}=\lambda a d^p$ with λ in K and $a^S a^{-1}=\beta^p$, then $a_{01}^S a_{01}^{-1}=\beta^p(d^S d^{-1})^p=\beta_0^p$. But then a_{01} is a constant multiple of a_1, and, by proper choice of λ, $a_1=a_{01}=\lambda a d^p$. The field $Z_0=K(z)$, $z=d^{-1}z_1$, $z^p=\lambda a$ is evidently equivalent to $K(z)$. Moreover $z^S=(d^S)^{-1}z_1^S=(d^S)^{-1}\beta d^S d^{-1}z=\beta z$ as desired.

We have determined the structure of cyclic fields of degree p^e over K when K contains a primitive pth root of unity ζ. We now study the more general case where ζ is not in the reference field F.

4. **The field $K=F(\zeta)$.** Let F be any field of characteristic not p so that the equation $x^p=1$ is separable and has as roots the primitive pth roots of unity

$$\zeta^i \qquad (i = 1, 2, \cdots, p-1), \tag{17}$$

---

* Such cyclic overfields define new quantities β_0 but we prove below that in fact we may replace β_0 by β.

and unity itself. Suppose that $h(x)$ is the irreducible factor in F of x^p-1 which has h as a root. Then the field $K=F(\zeta)$ is a normal field whose automorphisms form a group which is isomorphic to a subgroup of the cyclic group of order $p-1$ which replaces ζ by its powers (17). Every subgroup of a cyclic group is cyclic and hence K is cyclic of degree n over F. Moreover a generating automorphism of K over F is given by

$$T: \qquad \zeta \longleftrightarrow \zeta^t$$

where n divides $p-1$ and is prime to p, t is an integer *belonging* to the exponent n (mod p),

$$t^n \equiv 1 \pmod{p},\ t^e \not\equiv 1 \pmod{p},\ e < n. \tag{18}$$

If we define

$$\zeta_k = \zeta^{t_k},\ t_k \equiv t^{k-1} \pmod{p},\ 1 \leqq t_k < p, \tag{19}$$

$$\rho t \equiv 1 \pmod{p},\ \rho_k \equiv \rho^{k-1} \pmod{p}, \tag{20}$$

then I have proved*

LEMMA 1. *A quantity $\mu=\mu(\zeta)$ of I has the property*

$$\mu^T = \mu(\zeta^t) = \delta^p\mu^t \tag{21}$$

with δ in K if and only if there exists a quantity $\lambda=\lambda(\zeta)$ in K such that

$$\mu = \prod_{k=1}^{n}\lambda(\zeta_k)^{\rho_k}. \tag{22}$$

We shall also require the known*

LEMMA 2. *A cyclic field Z_0 of degree p over K, $Z_0=K(z)$, $z^p=\mu$ in K, is cyclic of degree pn over F, so that*

$$Z_0 = Z \times K, \tag{23}$$

where Z is cyclic of degree p over F, if and only if μ satisfies (21).

5. **Cyclic fields of degree p^e** over F. Let Z be cyclic of degree p^e over F. Then $Z_0=Z\times K$ is evidently cyclic of degree np^e over F and cyclic of degree p^e over K. Moreover Z contains a cyclic field Y of degree p^{e-1} over F and the field $Y_0=Y\times K$ is cyclic of degree np^{e-1} over F with automorphism group

$$S^iT^j \qquad (i=0,1,\cdots,p^{e-1}-1;\ j=0,1,\cdots,n-1).$$

By Theorem 2 we have

* Cf. *On normal Kummer fields*, etc., Lemma 3, Theorem 2.

THEOREM 4. *Let Z, Z_0, Y, Y_0 be defined as above. Then Y_0 contains a quantity β such that*

$$N_{Y_0/K}(\beta) = \zeta \tag{24}$$

and $Z_0 = Y_0(z)$, $z^p = \alpha$ in Y_0 such that

$$\alpha^S \alpha^{-1} = \beta_0{}^p. \tag{25}$$

Let a be a fixed quantity satisfying the equation (25) in α so that every solution α of (25) satisfies the condition

$$\alpha = \lambda a, \ \lambda \text{ in } K. \tag{26}$$

Then we have proved that z may always be chosen so that

$$z^S = \beta z, \tag{27}$$

for any β satisfying (24). We may then normalize the quantity β and prove

THEOREM 5. *The quantities β, a may be chosen so that*

$$\beta^T = \delta^p \beta^t, \ a^T = d^p a^t, \tag{28}$$

with δ, d in Y.

For we have $a^S = a\beta^p$ and may define

$$\beta_0 = \prod_{k=1}^{n} \beta(\zeta_k)^{\rho k}, \quad a_0 = \prod_{k=1}^{n} a(\zeta_k)^{\rho k}, \tag{29}$$

so that by Lemma 1 we have $\beta_0{}^T = \delta_0{}^p \beta_0{}^t$, $a_0{}^T = d_0{}^p a_0{}^t$. Since $ST = TS$ in Y, we also have

$$\begin{aligned} a_0^S a_0^{-1} &= \prod_{k=1}^{n} [a^S(\zeta_k)^{\rho k}][a(\zeta_k)^{\rho k}]^{-1} \\ &= \prod_{k=1}^{n} \beta(\zeta_k)^{\rho k \cdot p} = \beta_0^p. \end{aligned} \tag{30}$$

We also compute

$$N_{Y_0/K}(\beta_0) = \prod_{k=1}^{n} N_{Y_0/K} \beta(\zeta_k)^{\rho k} = \prod_{k=1}^{n} \zeta_k{}^{\rho k} = \zeta^\tau$$

where

$$\tau = \sum_{k=1}^{n} t_k \rho_k \equiv \sum_{k=1}^{n} (t\rho)^{k-1} \equiv n \pmod{p}. \tag{31}$$

Hence $N_{Y_0/K}(\beta_0) = \zeta^n$. We let $\mu n \equiv 1 \pmod{p}$, $\beta_1 = \beta_0{}^\mu$, $a_1 = a_0{}^\mu$ so that

$$N_{Y_0/K}(\beta_1) = \zeta^{\mu n} = \zeta, \tag{32}$$

and obviously

$$a_1^S a_1^{-1} = \beta_1^p. \tag{33}$$

Moreover

$$\beta_1^T = (\beta_0^T)^\mu = (\delta_0^p \beta_0^t)^\mu = (\delta_0^\mu)^p \beta_1^t = \delta^p \beta_1^t, \tag{34}$$

$$a_1^T = (a_0^T)^\mu = (d_0^p a_0^t)^\mu = (d_0^\mu)^p a_1^t = d^p a_1^t, \tag{35}$$

as desired. We have proved Theorem 5.

The automorphisms S and T of Y are commutative so that $N(\beta^T) = [N(\beta)]^T = \zeta^t = N(\beta^t)$ with $N(\beta)$ defined as $N_{Y_0/K}(\beta)$. Then by Theorem 1

$$\beta^T = f^S f^{-1} \beta^t \tag{36}$$

with f in Y_0. Also

$$\begin{aligned}(a^S a^{-1})^T &= (\beta^T)^p = a^{TS}(a^T)^{-1} = (d^S d^{-1})^p (a^S a^{-1})^t \\ &= (d^S d^{-1})^p \beta^{pt},\end{aligned} \tag{37}$$

so that

$$\beta^T = \zeta^\nu d^S d^{-1} \beta^t \qquad (0 \leqq \nu < p). \tag{38}$$

We shall only need (38) and $a^T = d^p a^t$ in our further study of the field Z.

We now take as basic in our study the given field $Y_0 = Y \times K$ of degree p^{e-1} over K where Y_0 is also cyclic of degree np^{e-1} over F and assume that Y_0 contains a quantity β such that $N_{Y_0/K}(\beta) = \zeta$. We have then shown that there always exists a quantity a of Y such that $a^S a^{-1} = \beta^p$ and moreover that β and a may be so chosen that (38) and

$$a^T = d^p a^t \qquad (d \text{ in } Y) \tag{39}$$

both hold. We now seek necessary and sufficient conditions that Y shall possess cyclic overfields of degree p^e over F. We shall in fact prove the fundamental result

THEOREM 6. *The field Y possesses cyclic overfields Z of degree p^e over F if and only if in* (38) $\nu = 0$. *Moreover every such field is determined by $Z_0 = Y_0(z)$, $z^p = \alpha$ in Y such that*

$$\alpha = \lambda a, \ \lambda^T = \sigma^p \lambda^t \tag{40}$$

with σ in K, where then $Z_0 = Z \times K$, Z_0 is cyclic of degree np^e over F.

For we may write $Y_0 = Y(\zeta)$ so that if Z is cyclic of degree p^e over F with

Y as sub-field then $Z_0 = Y_0(z)$, $z^p = \alpha = \lambda a$ with λ in K. Moreover Z is cyclic of degree p over Y and by Lemma 2 we have

$$\alpha^T = \psi^p \alpha^t \tag{41}$$

with ψ in Y. Hence

$$\lambda^T a^T = \lambda^T d^p a^t = \psi^p \lambda^t a^t, \tag{42}$$

and

$$\lambda^T = (\psi d^{-1})^p \lambda^t. \tag{43}$$

The quantity $x_1 = d^{-1}\psi$ has its pth power $x_1^p = \rho = \lambda^T \lambda^{-t}$ in K. Hence either $\psi = d\sigma$ with σ in K or $X_{10} = K(x_1)$ is a cyclic sub-field of Y_0 of degree p over K. But $Y_0 = Y \times K$ so that then $X_{10} = X_1 \times K$ where X is cyclic of degree p over F and in fact

$$\rho^T = \sigma^p \rho^t, \tag{44}$$

with γ in K. Then $\lambda^T = \lambda^t \rho$ implies

$$\begin{aligned} \lambda^{T^2} &= \lambda^{t^2} \rho^t \rho^T = \lambda^{t^2} \sigma_1^p \rho^{2t}, \\ \lambda^{T^3} &= \lambda^{t^3} \rho^{t^2} (\sigma^T)^p (\sigma^{2t})^p \rho^{2t^2} = \gamma_2^p \lambda^{t^3} \rho^{3t^2}, \end{aligned} \tag{45}$$

so that finally

$$\lambda^{T^n} = \lambda = \overset{p}{\gamma}_{n-1} \lambda^{t^n} \rho^{nt^{n-1}}. \tag{46}$$

The quantity $\lambda^{t^n - 1} = \lambda_0^p$ since $t^n \equiv 1 \pmod{p}$. Hence ρ^ϕ is the pth power of a quantity of K where $\phi = nt^{n-1}$ is prime to p. This evidently implies that ρ is the pth power of a quantity of K contrary to hypothesis. Hence $x_1 = \sigma$ in K and we have proved that (40) holds.

We have shown that z may be so chosen that $z^S = \beta z$ with (38), (39). Then (38) may be replaced by

$$\beta^T = \zeta^\nu (\psi^S \psi^{-1}) \beta^t, \tag{47}$$

since $\psi = \sigma d$, $\psi^S = \sigma d^S$.

Since $ST = TS$ in Z we obtain $(z^T)^p = \alpha^T = \psi^p \alpha^t = \psi^p z^{tp}$, $z^T = \zeta^\epsilon \psi z^t$ with $0 \leqq \epsilon < p$. Then $z^S = \beta z$ gives

$$z^{TS} = \zeta^\epsilon \psi^S \beta^t z^t = z^{ST} = (\beta z)^T = \zeta^\nu \psi^S \psi^{-1} \beta^t \zeta^\epsilon \psi z^t, \tag{48}$$

so that $\zeta^\nu = 1$, $\nu = 0$.

Conversely let Y be cyclic of degree p^{e-1} over F, $Y_0 = Y \times K$, β and a be chosen in Y_0 and satisfying $N_{Y_0/K}(\beta) = \zeta$, (38), (39). Let λ range over all quantities of K such that (40) holds so that α satisfies (47). We have proved

that then $Z_0=Y_0(z)$ has the property $Z_0=K(z)$ and is cyclic of degree p^e over K. It remains merely to show that then Z_0 is actually cyclic of degree $p^e n$ over F if $\nu=0$. We define the automorphism T of Z_0 by that in Y_0 and by

$$z^T = \psi z^t, \ \psi = \sigma d,$$

where $\alpha^T=\psi^p\alpha^t$. Then we require only to show that $ST=TS$ so that the automorphism group of Z_0 over F is actually the cyclic group (S^iT^j) $(i=0, 1, \cdots, p^e-1; j=0, 1, \cdots, n-1)$. But this immediately follows from the computation in (48) with $\epsilon=0$, and Theorem 6 is proved.

University of Chicago,
Chicago, Ill.

A CANONICAL POWER SERIES EXPANSION FOR A SURFACE*

BY

ERNEST P. LANE

1. **Introduction.** In studying the projective differential geometry of an analytic surface in ordinary space it is frequently convenient to use a power series expansion for one non-homogeneous projective coordinate of a variable point on the surface in terms of the other two coordinates. Such an expansion can be reduced to various canonical forms by means of suitable choices of the projective coordinate system, that is, of the tetrahedron of reference and the unit point. The purpose of this paper is to deduce and apply such a canonical expansion, which does not seem to have been considered hitherto.

The form of the canonical expansion with which we are concerned was suggested in the first place by the exigencies adherent to the investigation of problems of a certain type which will be discussed more fully in §3. For the moment it is sufficient to say that the desideratum was a canonical expansion for which one edge of the tetrahedron of reference was an arbitrary non-asymptotic tangent at an ordinary non-parabolic point of a surface. This consideration led to the deduction of the canonical expansion which will be explained in detail in §2. There the method followed is predominantly geometrical and the coordinate system employed is completely interpreted geometrically.

In §4 the connection between the canonical expansion introduced in this paper and a canonical expansion long known in the theory of conjugate nets is established, and some of the results of previous sections are applied in the theory of conjugate nets.

2. **The canonical expansion.** In this section a canonical power series expansion for a surface will be deduced. The geometrical significance of the successive steps in the work will be fully explained.

In projective space of three dimensions let us establish a point coordinate system, in which a point that has non-homogeneous coordinates x, y, z also has homogeneous coordinates x_1, x_2, x_3, x_4 whose ratios are defined by placing

$$x = \frac{x_2}{x_1}, \qquad y = \frac{x_3}{x_1}, \qquad z = \frac{x_4}{x_1}.$$

* Presented to the Society, April 20, 1935; received by the editors December 5, 1934.

Then let us consider an analytic surface S whose equation is

$$z = f(x, y). \tag{1}$$

It is well known that if the origin O, namely, the vertex of the tetrahedron of reference having non-homogeneous coordinates 0, 0, 0, is an ordinary point on the surface S, and if the face $z=0$ of the tetrahedron is the tangent plane of S at O, then the equation (1) of the surface can be expanded by Taylor's series into

$$z = \alpha(x, y) + \beta(x, y) + \gamma(x, y) + \cdots, \tag{2}$$

where $\alpha, \beta, \gamma, \cdots$ are forms which can be written as follows:

$$\begin{aligned} \alpha(x, y) &= a_0x^2 + 2a_1xy + a_2y^2, \\ \beta(x, y) &= b_0x^3 + 3b_1x^2y + 3b_2xy^2 + b_3y^3, \\ \gamma(x, y) &= c_0x^4 + 4c_1x^3y + 6c_2x^2y^2 + 4c_3xy^3 + c_4y^4, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \end{aligned} \tag{3}$$

the coefficients $a_0, \cdots, c_4, \cdots$ being constants. Let us assume that the origin O is not a parabolic point. Then the discriminant Δ of the form α, defined by

$$\Delta = 4(a_0a_2 - a_1^2), \tag{4}$$

is not zero, and the two asymptotic tangents at the origin, represented by the equation $\alpha=0$, are distinct.

The principal purpose of this section is to prove the following theorem:

By means of a suitable choice of the coordinate system, the equation (2) *of a surface can be reduced to the canonical form*

$$z = x^2 + y^2 + x^3 + b_3y^3 + c_0x^4 + c_4y^4 + \cdots, \tag{5}$$

the unwritten terms being of degree at least five, and the edge $y=z=0$ of the tetrahedron of reference being an arbitrary non-asymptotic tangent at the non-parabolic point (0, 0, 0).

The proof begins with the observation that the coefficient a_1 can be made to vanish by taking for the edge $x=z=0$ of the tetrahedron of reference the tangent which is the harmonic conjugate of the edge $y=z=0$ with respect to the asymptotic tangents of the surface at the origin O. We shall suppose from now on that this has been done, so that $a_1=0$.

The next step in the proof consists in establishing the following lemma:

The coefficient b_1 vanishes if, and only if, the vertex (0, 0, 1, 0) *corresponds to the face $y=0$ in Segre's correspondence, and similarly b_2 vanishes in case the vertex* (0, 1, 0, 0) *corresponds to the face $x=0$.*

Let us recall in this connection that *Segre's correspondence* associated with a point O of a surface S is by definition the correspondence between the *osculating plane* at O of any curve C on S through O and the *ray-point* of the curve C at O, namely, the point of intersection of three tangent planes of S constructed at O and at two consecutive points on C. Segre studied* this correspondence, using a notation not differing essentially from ours, and proved that the equations of the correspondence can be written in the form

$$(6) \qquad X = \frac{\alpha(u, v)\alpha_v(u, v)}{\Delta\alpha(u, v) - J(u, v)}, \qquad Y = -\frac{\alpha(u, v)\alpha_u(u, v)}{\Delta\alpha(u, v) - J(u, v)},$$

the meanings of the new symbols appearing here having the following explanation. The symbols X, Y represent the first two non-homogeneous coordinates of a point in the tangent plane of the surface S at the origin O, the third coordinate being zero. Then u, v are coordinates of a plane through O, the equation of this plane being

$$(7) \qquad z = vx - uy.$$

Subscripts indicate partial differentiation, and J is the jacobian of the forms α, β:

$$(8) \qquad J = \alpha_u\beta_v - \alpha_v\beta_u.$$

Writing out equations (6) by means of equations (3) with $a_1 = 0$, placing $v = 0$, and then allowing u to become infinite, we find that the homogeneous coordinates of the point corresponding to the plane $y = 0$ are $3b_1$, 0, a_0, 0. Similarly, placing $u = 0$ and then letting v become infinite, we find that the point $(3b_2, a_2, 0, 0)$ corresponds to the plane $x = 0$. One sees immediately that the first point is the vertex $(0, 0, 1, 0)$ of the tetrahedron of reference if, and only if, $b_1 = 0$; similarly, the second point is the vertex $(0, 1, 0, 0)$ in case $b_2 = 0$. Thus the lemma is established. We shall suppose from now on that $b_1 = b_2 = 0$.

It is geometrically obvious that the face $x = 0$ may still be any plane whatever through the tangent $x = z = 0$, except the tangent plane, and likewise the face $y = 0$ is an arbitrary plane through the tangent $y = z = 0$. This fact can also be verified analytically in the following way. Let us make the transformation

$$(9) \qquad \begin{gathered} x = \frac{x' + Bz'}{1 + 2a_0Bx' + 2a_2Dy'}, \qquad y = \frac{y' + Dz'}{1 + 2a_0Bx' + 2a_2Dy'}, \\ z = \frac{z'}{1 + 2a_0Bx' + 2a_2Dy'}, \end{gathered}$$

* Segre, *Complementi alla teoria delle tangenti coniugate di una superficie*, Rendiconti della Reale Accademia dei Lincei, (5), vol. 17 (1908), p. 405.

from the old coordinates x, y, z to new coordinates x', y', z', the coefficients B, D being arbitrary. This transformation makes the two planes whose equations in the old coordinate system are

$$x - Bz = 0, \qquad y - Dz = 0 \tag{10}$$

become, respectively, the two faces $x'=0$ and $y'=0$ of the tetrahedron of reference of the new coordinate system. Moreover, this transformation carries the equation (2) of the surface, with $a_1 = b_1 = b_2 = 0$, into another equation of the same form in which the coefficients of the terms of the second and third degrees are absolutely unchanged. The calculations in support of this statement will not be reproduced here, but are straightforward, and complete the analytic verification aforementioned.

The transformation (9) can be used to reduce both of c_1, c_3 to zero. If the transformed coefficients of the terms of the fourth degree are calculated, it is found that, in particular, the coefficients c_1, c_3 obey the laws

$$c_1' = c_1 - \tfrac{1}{2}a_2b_0D, \qquad c_3' = c_3 - \tfrac{1}{2}a_0b_3B. \tag{11}$$

Therefore, if B, D have the values

$$B = \frac{2c_3}{a_0b_3}, \qquad D = \frac{2c_1}{a_2b_0}, \tag{12}$$

then the transformed coefficients c_1', c_3' are zero. The equations in the old coordinate system of the planes which must be taken as the faces $x'=0$ and $y'=0$ of the new tetrahedron in order to effect this reduction are found, by substituting the values of B, D from (12) into (10), to be

$$a_0b_3x - 2c_3z = 0, \qquad a_2b_0y - 2c_1z = 0. \tag{13}$$

A geometric characterization of these planes is contained in the following lemma:

The first of the planes (13) *is the plane that passes through the tangent line* $x=z=0$ *and touches the cone of Segre associated with the point O of the surface; likewise the second of the planes* (13) *is the plane through the tangent line* $y=z=0$ *and touching the cone of Segre.*

Beginning the proof, let us recall that *the cone of Segre* at a point O of a surface S is by definition the envelope of the planes through O which are able to serve as stationary osculating planes at O of curves on S through O and having at O stationary ray-points. The equation of the cone of Segre, when the equation of the surface is in the general form (2), was shown* by Segre to be

* Segre, ibid.

$$\tfrac{1}{2}\alpha(\beta_{uu}\alpha_v - 2\beta_{uv}\alpha_u\alpha_v + \beta_{vv}\alpha_u^2 - 2\Delta\beta) + 2\beta J - \alpha K = 0, \tag{14}$$

in which the arguments of the functions are the plane coordinates u, v, and the symbols Δ, J have the same meanings as in equations (6), while K is the jacobian of the forms α, γ:

$$K = \alpha_u\gamma_v - \alpha_v\gamma_u. \tag{15}$$

To complete the proof it is sufficient to write out equation (14) by means of (3) with $a_1 = b_1 = b_3 = 0$, and then, on comparing each of equations (13) with equation (7), to show that the equation of the cone of Segre is satisfied by the values of u, v given by

$$u = 0, \qquad v = \frac{a_0 b_3}{2c_3},$$

and is also satisfied by

$$u = -\frac{a_2 b_0}{2c_1}, \qquad v = 0.$$

We shall suppose from now on that $c_1 = c_3 = 0$.

The position of the vertex (0, 0, 0, 1) on the edge $x = y = 0$ of the tetrahedron of reference is still at our disposal. It will now be shown that this vertex can be chosen so as to reduce c_2 to zero. In fact, the transformation of coordinates

$$x = \frac{x'}{1 + Hz'}, \qquad y = \frac{y'}{1 + Hz'}, \qquad z = \frac{z'}{1 + Hz'}, \tag{16}$$

in which the coefficient H is arbitrary, makes the point whose coordinates in the old coordinate system are H, 0, 0, 1 become the vertex (0, 0, 0, 1) of the tetrahedron of reference of the new coordinate system. Moreover, this transformation changes the equation of the surface into another equation of the same form in which the coefficients of the terms of the second and third degrees are absolutely unchanged, and in which the coefficients c_1' and c_3' are still zero, while the coefficient c_2' is given by the formula

$$c_2' = c_2 - \tfrac{1}{3}a_0 a_2 H. \tag{17}$$

Therefore, if H has the value given by

$$H = \frac{3c_2}{a_0 a_2}, \tag{18}$$

then $c_2' = 0$. The homogeneous coordinates in the old coordinate system of the

point V which must be taken as the vertex $(0, 0, 0, 1)$ of the tetrahedron of reference of the new coordinate system in order to effect this reduction are

$$\frac{3c_2}{a_0a_2},\ 0,\ 0,\ 1. \tag{19}$$

A simple geometric construction for the point V is certainly desirable. The following discussion shows how to construct this point by means of its relations to certain other points on the line $x=y=0$, namely, the points distinct from the origin O, in which this line intersects certain quadrics associated with the point O of the surface S.

First of all, the equation of *the canonical quadric of Wilczynski* can be found by means of the characteristic property discovered* by Bompiani to the effect that this quadric is an asymptotic osculating quadric of a curve on a surface which has an inflexion at the point under consideration. The details of the calculation of the equation need not be reproduced here, but the required *equation of the canonical quadric of Wilczynski* turns out to be

$$z = a_0x^2 + a_2y^2 + \frac{3}{4}\frac{b_0}{a_0}xz + \frac{3}{4}\frac{b_3}{a_2}yz + Wz^2, \tag{20}$$

the coefficient W being given by the formula

$$W = \frac{3}{8}\left[\frac{c_0}{a_0^2} + \frac{2c_2}{a_0a_2} + \frac{c_4}{a_2^2} - \frac{3}{4}\left(\frac{b_0^2}{a_0^3} + \frac{b_3^2}{a_2^3}\right)\right]. \tag{21}$$

The equation of *the quadric of Lie* can be found by means of its definition as the quadric determined by three consecutive asymptotic tangents of one family constructed at the point O under consideration and at two consecutive points on the asymptotic curve of the other family through O. The required *equation of the quadric of Lie* has the same form as equation (20) except that the coefficient W is replaced by a coefficient L given by

$$L = \frac{3}{8}\left[\frac{c_0}{a_0^2} + \frac{2c_2}{a_0a_2} + \frac{c_4}{a_2^2} - \frac{15}{16}\left(\frac{b_0^2}{a_0^2} + \frac{b_3^2}{a_2^3}\right)\right]. \tag{22}$$

As a matter of fact, the equation (20) with W arbitrary represents any quadric of Darboux.

The equation of an *asymptotic osculating quadric* at the point O for the curve cut on the surface by the plane $y=0$ can be found from its definition as the quadric determined by three consecutive asymptotic tangents of one

* Bompiani, *Fascio di quadriche di Darboux e normale proiettiva in un punto di una superficie*, Rendiconti della Reale Accademia dei Lincei, (6), vol. 6 (1927), p. 187.

family constructed at O and two consecutive points on the curve. We find that *the equations of the two asymptotic osculating quadrics at the point O for the curve* $y=0$ can be written together in the form

$$z = a_0x^2 + a_2y^2 + \frac{9}{4}\frac{b_0}{a_0}xz \mp \frac{15}{4}\frac{b_0}{a_0w}yz + Bz^2, \tag{23}$$

where w is defined by

$$w = \left(-\frac{a_0}{a_2}\right)^{1/2}, \tag{24}$$

and the coefficient B is given by

$$B = \frac{3}{2}\left(\frac{c_0}{a_0^2} - \frac{c_2}{a_0a_2} - \frac{9}{8}\frac{b_0^2}{a_0^3}\right). \tag{25}$$

The analogous equation of the two asymptotic osculating quadrics at the point O for the curve $x=0$ can be written by symmetry and is

$$z = a_0x^2 + a_2y^2 + \frac{9}{4}\frac{b_3}{a_2}yz \mp \frac{15}{4}\frac{b_3w}{a_2}xz + Kz^2, \tag{26}$$

where the coefficient K is given by

$$K = \frac{3}{2}\left(\frac{c_4}{a_2^2} - \frac{c_2}{a_0a_2} - \frac{9}{8}\frac{b_3^2}{a_2^3}\right). \tag{27}$$

The coordinates of the point distinct from O in which the line $x=y=0$ pierces the canonical quadric of Wilczynski are W, 0, 0, 1. We may speak of this point as *the point W*, since W is a non-homogeneous coordinate of this point on the line $x=y=0$. With similar terminology in analogous cases, the following construction locates the point H which we are seeking to characterize.

Let P be the harmonic conjugate of the point W with respect to the points L and O. Let Q be the harmonic conjugate of O with respect to B and K. Then the point H is the harmonic conjugate of Q with respect to P and O.

The proof consists in showing first that the coordinates P, Q of the points P, Q are given by the formulas

$$\begin{aligned} P &= \frac{3}{8}\left[\frac{c_0}{a_0^2} + \frac{2c_2}{a_0a_2} + \frac{c_4}{a_2^2} - \frac{9}{8}\left(\frac{b_0^2}{a_0^3} + \frac{b_3^2}{a_2^3}\right)\right], \\ Q &= \frac{3}{4}\left[\frac{c_0}{a_0^2} - \frac{2c_2}{a_0a_2} + \frac{c_4}{a_2^2} - \frac{9}{8}\left(\frac{b_0^2}{a_0^3} + \frac{b_3^2}{a_2^3}\right)\right], \end{aligned} \tag{28}$$

and in then showing that the point H, given by equation (18), is such that the cross ratio of the points O, P, Q, H in the order written is -1.

The vertices of the tetrahedron of reference are now completely specified when the point O and the tangent $y=z=0$ are given. We shall suppose from now on that $c_2=0$, and have shown how to reduce the equation (2) of a surface to the form

$$z = a_0x^2 + a_2y^2 + b_0x^3 + b_3y^3 + c_0x^4 + c_4y^4 + \cdots, \tag{29}$$

the unwritten terms being of degree at least five.

There remains the transformation of unit point,

$$x = lx', \qquad y = my', \qquad z = nz', \tag{30}$$

the coefficients l, m, n being arbitrary. If this transformation is effected on equation (29), if the transformed equation is solved for z' and the coefficients a_0', a_2', b_0' are equated to unity, the following conditions on l, m, n are obtained:

$$a_0l^2 = a_2m^2 = b_0l^3 = n. \tag{31}$$

Solving these equations for l, m, n, we find

$$l = \frac{a_0}{b_0}, \qquad m = \frac{a_0}{b_0}\left(\frac{a_0}{a_2}\right)^{1/2}, \qquad n = \frac{a_0^3}{b_0^2}, \tag{32}$$

and thus arrive at the following lemma:

The unit point can be chosen so as to make $a_0=a_2=b_0=1$.

The theorem stated at the beginning of this section has now been proved. However, it still remains to describe geometrically the point whose coordinates in the old coordinate system are

$$\frac{a_0}{b_0}, \qquad \frac{a_0}{b_0}\left(\frac{a_0}{a_2}\right)^{1/2}, \qquad \frac{a_0^3}{b_0^2}, \tag{33}$$

which becomes the unit point in the new coordinate system. This point will be located as the intersection of three planes.

First of all, this point obviously lies on one of the two planes whose equation is

$$a_0x^2 - a_2y^2 = 0. \tag{34}$$

These are the planes determined by the line $x=y=0$ and *the associate conjugate tangents* of the tangents $x=z=0$ and $y=z=0$, that is, the tangents which separate the latter harmonically and also separate the asymptotic tangents harmonically.

To find a second plane containing the point (33) let us observe that the equation of *the quadric of Moutard* for the tangent $y=z=0$ is

$$z = a_0x^2 + a_2y^2 + \frac{b_0}{a_0}xz + \left(\frac{c_0}{a_0^2} - \frac{b_0^2}{a_0^3}\right)z^2, \tag{35}$$

as can be verified by finding the locus of the osculating conics at the point O of the curves of section of the surface made by planes through this tangent. The equation of the tangent plane of this quadric at the point (0, 1, 0, 0) is

$$2a_0^2x + b_0z = 0. \tag{36}$$

The harmonic conjugate of the plane $x=0$ with respect to the tangent plane, $z=0$, and the plane (36) has the equation

$$a_0^2x + b_0z = 0. \tag{37}$$

Finally, the harmonic conjugate of the plane (37) with respect to the planes $z=0$ and $x=0$ is represented by the equation

$$a_0^2x - b_0z = 0, \tag{38}$$

and obviously contains the point (33).

A third plane containing this point can be found in the following way. The cubic surface whose equation is

$$z = a_0x^2 + a_2y^2 + b_0x^3 + b_3y^3 \tag{39}$$

is completely characterized by the following properties: it has third-order contact with the surface at the point O; it has a unode at the point (0, 0, 0, 1); and its uniplane is the plane $x_1=0$. Now the polar plane of the point (0, 1, 0, 0) with respect to the cubic (39) has the equation

$$a_0x_1 + 3b_0x_2 = 0 \tag{40}$$

in homogeneous coordinates. The equation of the harmonic conjugate of this plane with respect to the planes $x_1=0$ and $x_2=0$ is

$$a_0x_1 - 3b_0x_2 = 0. \tag{41}$$

The equation of the harmonic conjugate of the plane (41) with respect to the plane $x_1=0$ and the plane (40) is

$$a_0x_1 + b_0x_2 = 0. \tag{42}$$

Finally, the harmonic conjugate of the plane (42) with respect to the planes $x_1=0$ and $x_2=0$ has the equation

$$a_0x_1 - b_0x_2 = 0, \tag{43}$$

and evidently contains the point (33), whose geometric description is now complete.

Two remarks may be adjoined. Since the derivative $\partial^2 z/\partial x\partial y$ calculated from equation (39) vanishes identically, it follows that *the pencils of planes whose axes are the lines* $x=x_1=0$ *and* $y=x_1=0$ *cut on the cubic surface* (39) *a conjugate net.* In the second place, a similar remark can be made about the quadric surface whose equation is

$$z = a_0x^2 + a_2y^2, \tag{44}$$

and which is completely characterized by the following properties. It has second-order contact with the surface at the point O; the lines $z=x_1=0$ and $x=y=0$ are reciprocal polars with respect to it; and it passes through the point $(0, 0, 0, 1)$.

3. **Plane sections through a tangent.** The canonical expansion deduced in the preceding section was designed to be used in solving problems of a certain type which may be identified as follows. Let us pass a plane through a non-asymptotic tangent at a point O of a surface and consider geometric elements projectively associated with the point O of the plane curve of intersection of the surface and plane. The problem is to discuss the loci of these elements when the plane turns around the tangent line.

Moutard was probably the first to solve a projective problem of this type when he proved that the locus of the osculating conics of the plane curves of section made by planes through a tangent is a quadric surface, which now bears his name. Kubota and Su have also solved* projective problems of this type, by showing among other things that the locus of the projective normals is a cubic cone of the third order. The purpose of this section is to continue and extend the investigations of Kubota and Su. Since the projective normal at a point of a plane curve is a nodal tangent of the eight-point nodal cubic, it is proposed here to consider all the eight-point cubics, and the osculating cubic, of a plane curve of section of a surface, and to investigate their loci and the loci of various points and lines associated with them, when the plane turns around the tangent.

Let us consider a surface S whose equation in non-homogeneous projective coordinates has been reduced to the canonical form

$$\begin{aligned} z = {} & x^2 + y^2 + x^3 + Ay^3 + Bx^4 + Cy^4 + Dx^5 + Ex^4y + Fx^3y^2 \\ & + Gx^2y^3 + \cdots + Hx^6 + Ix^5y + Jx^4y^2 + \cdots + Kx^7 + Lx^6y \\ & + \cdots + Mx^8 + \cdots , \end{aligned} \tag{45}$$

* Kubota and Su, *Eine Bemerkung zur Projektivdifferentialgeometrie der Flächen*, Science Reports of the Tôhoku Imperial University, (1), vol. 19 (1930), p. 293.

all the coefficients of which are absolute invariants of the surface, the terms written being the only ones that will be needed hereinafter. The equation

$$y = nz \qquad (n \neq 0) \tag{46}$$

represents a plane π through the line t whose equations are $y=z=0$ and which is an ordinary tangent of the surface S at the origin O (0, 0, 0). The plane π cuts through the surface S in a curve whose projection from the point (0, 0, 0, 1) onto the tangent plane, $z=0$, of S at O is a curve C. The equation, in the tangent plane, of the curve C is found, by eliminating z between equations (45), (46), to be

$$\begin{aligned}\frac{y}{n} &= x^2 + y^2 + x^3 + Ay^3 + Bx^4 + Cy^4 + Dx^5 + Ex^4y + Fx^3y^2\\ &\quad + Gx^2y^3 + \cdots + Hx^6 + Ix^5y + Jx^4y^2 + \cdots + Kx^7 + Lx^6y\\ &\quad + \cdots + Mx^8 + \cdots .\end{aligned} \tag{47}$$

If this equation is solved for y as a power series in x, the result to terms of the eighth degree is

$$y = n(x^2 + x^3 + b_4x^4 + b_5x^5 + b_6x^6 + b_7x^7 + b_8x^8 + \cdots), \tag{48}$$

the coefficients $b_4, \cdots, b_8$ being defined by the formulas

$$\begin{aligned} b_4 &= n^2 + B,\\ b_5 &= 2n^2 + D,\\ b_6 &= 2n^4 + An^3 + (2B+1)n^2 + En + H,\\ b_7 &= 6n^4 + 3An^3 + (2B + 2D + F)n^2 + (E+I)n + K,\\ b_8 &= 5n^6 + 5An^5 + (6B + C + 6)n^4 + (3A + 3E + G + 3AB)n^3\\ &\quad + (2D + 2F + 2H + J + B^2)n^2 + (I + L + BE)n + M.\end{aligned} \tag{49}$$

The equation of *the osculating conic* of the curve C at the origin O is found by writing the most general equation of the second degree in x, y and then demanding that this equation be satisfied by the series (48) for y identically in x as far as the term in x^4. The result is

$$y - nx^2 - xy - R\frac{y^2}{n} = 0, \tag{50}$$

where R is a polynomial defined by placing

$$R = n^2 + B - 1. \tag{51}$$

The equation of *the quadric of Moutard* for the tangent t at the point O is found, by eliminating n between equations (46), (50), to be

$$z - x^2 - y^2 - xz - (B-1)z^2 = 0. \tag{52}$$

Let the left member of equation (50) be denoted by Q, so that

$$Q = y - nx^2 - xy - R\frac{y^2}{n}. \tag{53}$$

Then substitution of the series (48) for y yields, to terms of the fifth degree,

$$Q = Snx^5 + \cdots, \tag{54}$$

the coefficient S being defined by the formula

$$S = -n^2 - 3B + D + 2. \tag{55}$$

The two values of n which make S vanish yield, when substituted in equation (46), the equations of the two planes producing sections of the surface which are hyperosculated by their osculating conics. These equations can be written together in the form.

$$y^2 - (-3B + D + 2)z^2 = 0, \tag{56}$$

which establishes the following theorem:

At a point O of a surface S the two planes through a tangent t which produce sections which are hyperosculated by their osculating conics separate harmonically the tangent plane and the plane through the tangent t which touches the cone of Segre.

The equation of *all eight-point cubics* at the point O of the curve C is found by writing the most general equation of the third degree in x, y and then demanding that this equation be satisfied by the series (48) for y identically in x as far as the term in x^7. The result can be written in the form

$$\begin{aligned} S\left[y - nx^2 - nx^3 - (R+1)x^2y - (S+2R)x\frac{y^2}{n} - (V+2S+R^2)\frac{y^3}{n^2}\right] \\ - WQ\frac{y}{n} + h\left[\left(Sx - V\frac{y}{n}\right)Q - S^2\frac{y^3}{n^2}\right] = 0, \end{aligned} \tag{57}$$

in which h is a parameter and V, W are two polynomials defined by the formulas

$$\begin{aligned} V &= An^3 - (2B-3)n^2 + En + 10B - 4D + H - 2B^2 - 5, \\ W &= -n^4 - (D - F - 2)n^2 - (2E - I)n - 5B + 5D - 3H + K \\ &\quad - 3BD + 5B^2. \end{aligned} \tag{58}$$

The equation of *the eight-point nodal cubic* is

$$\left(Sx - V\frac{y}{n}\right)Q - S^2\frac{y^3}{n^2} = 0. \tag{59}$$

The nodal tangents of this cubic are *the tangent t*, whose equation is $y=0$, and *the projective normal*, whose equation is

$$Sx - V\frac{y}{n} = 0. \tag{60}$$

The equation of *the flex-ray*, that is, the line containing the three inflexions of the eight-point nodal cubic, is found to be

$$S^2 - (2VS + S^2)x + (V^2 - RS^2)\frac{y}{n} = 0. \tag{61}$$

The flex-ray is known to be tangent to the osculating conic (50) at *the point of intersection of the flex-ray* (61) *and the projective normal* (60), namely, the point

$$\left(\frac{VS}{V^2 + VS + RS^2}, \frac{S^2 n}{V^2 + VS + RS^2}\right). \tag{62}$$

Moreover, the flex-ray intersects the tangent t at the point

$$\left(\frac{S}{2V + S}, 0\right). \tag{63}$$

All the eight-point cubics (57) intersect, besides at the origin O, also in *the Halphen point*,

$$\begin{aligned}\Bigg(&\frac{ST(VT - S^4)}{(VT - S^4)^2 + ST(VT - S^4) + T^2(RS^2 - T)},\\ &\qquad\qquad \frac{S^2T^2 n}{(VT - S^4)^2 + ST(VT - S^4) + T^2(RS^2 - T)}\Bigg),\end{aligned} \tag{64}$$

wherein T is a polynomial defined by the formula

$$T = V^2 + 2VS + 2RS^2 - SW. \tag{65}$$

The equation of *the line joining the origin to the Halphen point is*

$$STx - (VT - S^4)\frac{y}{n} = 0. \tag{66}$$

The eight-point cubic of Sannia is by definition that eight-point cubic which passes through the point (63), or equally well, that eight-point cubic

which passes through the point (62). Either way one finds that for the eight-point cubic of Sannia the parameter h in equation (57) has the value

$$h = -2\left(\frac{V}{S} + 1\right). \tag{67}$$

With this value of h in equation (57), the equation of the eight-point cubic of Sannia can be written in the form

$$\begin{aligned} &\left[S^2 - (2VS + S^2)x + (V^2 - RS^2)\frac{y}{n} + T\frac{y}{n}\right]Q \\ &\qquad - S^2\left(Sx - V\frac{y}{n}\right)\frac{y^2}{n} = 0. \end{aligned} \tag{68}$$

Moreover, it is evident on inspection that equation (53) defining Q is equivalent to

$$S^2Q = \left[S^2 - (2VS + S^2)x + (V^2 - RS^2)\frac{y}{n}\right]y - n\left(Sx - V\frac{y}{n}\right)^2.$$

Therefore the flex-ray (61) meets the eight-point cubic of Sannia, besides in the points (62), (63), also in *the point where the flex-ray intersects the line* (66) *joining the origin to the Halphen point*, namely in the point

$$\begin{aligned} &\left(\frac{S(VT - S^4)}{(VT - S^4)(2V + S) - T(V^2 - RS^2)},\right. \\ &\qquad\qquad \left.\frac{S^2Tn}{(VT - S^4)(2V + S) - T(V^2 - RS^2)}\right). \end{aligned} \tag{69}$$

The osculating cubic is represented by equation (57) when the parameter h has the value obtained by demanding that (57) be satisfied by the series (48) for y as far as the term in x^8. Thus *the value of h for the osculating cubic* is found to be given by

$$\begin{aligned} Th = S[&An^5 - (B - C + 1)n^4 - (E - G + AB)n^3 \\ &+ (2F + J - 2H + 2B^2 + 3)n^2 - (3E - I - L + 3BE)n \\ &- 6B + 6D - 3H + M + 6BD - 4BH - B^2 - 2D^2 + 3B^3] \\ &- (V + 4S)W. \end{aligned} \tag{70}$$

The right member of this equation reduces to a polynomial of the sixth degree in n, as does T also.

The osculating conic (50) intersects any eight-point cubic (57), besides at the origin also at the point

$$\left(\frac{-S(V+2S+hS)}{(V+2S+hS)^2-S(V+2S+hS)+RS^2},\right.$$
$$\left.\frac{S^2n}{(V+2S+hS)^2-S(V+2S+hS)+RS^2}\right). \tag{71}$$

For the value of h given by (67) this point becomes the point (62), and for the value of h given by (70) this point is *the residual intersection of the osculating conic and the osculating cubic*. The equation of *the line joining the origin to the point* (71) *is obviously*

$$Sx+(V+2S+hS)\frac{y}{n}=0. \tag{72}$$

It is now easy to write the equation of *the cone of Kubota and Su*, which is *the locus of the projective normals*. Elimination of n between equations (46), (60) yields the desired equation,

$$\begin{aligned}&[y^2-(-3B+D+2)z^2]x+Ay^3+(-2B+3)y^2z+Eyz^2\\&\quad+(10B-4D+H-2B^2-5)z^3=0.\end{aligned} \tag{73}$$

The form of this equation makes it evident at once that the tangent t, $y=z=0$, is a double line of this cone, as remarked* by Kubota and Su. Moreover, the equation of the nodal tangent planes of the cone along this line is obtained by setting equal to zero the coefficient of x in equation (73), and is precisely equation (56). Thus the following theorem is proved.

The nodal tangent planes of the cone of Kubota and Su along its double line are the planes through this line that cut the surface in curves which are hyperosculated by their osculating conics.

Furthermore, the cone of Kubota and Su cuts the tangent plane, $z=0$, in the line

$$x+Ay=0, \tag{74}$$

besides the tangent t. Finally, the plane containing the three inflexional generators of the cone is represented by the equation

$$\begin{aligned}&4(3B-D-2)x+[E+3A(3B-D-2)]y\\&\quad+[(2B-3)(2-3B+D)+3(10B-4D+H-2B^2-5)]z=0.\end{aligned} \tag{75}$$

For the purpose of facilitating calculation it is convenient to introduce seven constants a, b, c, d, e, f, g defined by the following formulas:

* Kubota and Su, ibid., p. 300.

$$
\begin{aligned}
a &= B - 1, \\
b &= -3B + D + 2, \\
c &= -2B + 3, \\
d &= 10B - 4D + H - 2B^2 - 5, \\
e &= -D + F + 2, \\
f &= -2E + I, \\
g &= -5B + 5D - 3H + K - 3BD + 5B^2.
\end{aligned} \tag{76}
$$

Then one can rewrite quite simply the definitions of R, S, V, W:

$$
\begin{aligned}
R &= n^2 + a, \\
S &= -n^2 + b, \\
V &= An^3 + cn^2 + En + d, \\
W &= -n^4 + en^2 + fn + g.
\end{aligned} \tag{77}
$$

Actual calculation now yields

$$T = (1 + A^2)n^6 - 4A(B - 1)n^5 + \alpha n^4 + \beta n^3 + \gamma n^2 + \delta n + \epsilon,$$

where $\alpha, \beta, \gamma, \delta, \epsilon$ are defined by

$$
\begin{aligned}
\alpha &= 2a - 3b - 2c + e + c^2 + 2AE, \\
\beta &= 2(Ab + Ad + Ec - E) + f, \\
\gamma &= 2bc + 2cd - 2d + 2b^2 - 4ab - be + g + E^2, \\
\delta &= 2E(b + d) - bf, \\
\epsilon &= 2bd + 2ab^2 - bg + d^2.
\end{aligned}
$$

The values of n which satisfy the equation $T=0$ make the Halphen point (64) coincide with the origin, which is then a *coincidence point* on the corresponding curve of section. Thus one arrives at the following theorem:

Through an ordinary tangent at a point O of a surface there pass six planes which cut the surface in curves having a coincidence point at O.

It is known that at a coincidence point the eight-point nodal cubic is the osculating cubic. It is suggested to inquire whether the eight-point cubic of Sannia can be the osculating cubic. Demanding that equation (68) be satisfied by the power series (48) for y as far as the term in x^8, we find that n must be a solution of an equation of the ninth degree,

$$A(1 + A^2)n^9 + \cdots + 9 = 0, \tag{78}$$

the unwritten terms being not needed for our purposes. Thus the following theorem is established.

Through an ordinary tangent at a point O of a surface there pass nine planes which cut the surface in curves each of which is osculated at O by its eight-point cubic of Sannia.

Besides the projective normal there are other lines whose loci are of interest. For example, *the locus of the flex-ray* is an algebraic ruled surface of the sixth order, as remarked* by Kubota and Su, whose equation is found, by eliminating n from (61), (46), to be

$$\begin{aligned} &z(y^2 - bz^2)^2 - x[-2Ay^5 + (4B-5)y^4z + 2(Ab-E)y^3z^2 \\ &+ 2(bc - b - d)y^2z^3 + 2Ebyz^4 + (2bd + b^2)z^5] + (A^2-1)y^6 \\ &+ 2A(3-2B)y^5z + (2AE + c^2 + 2b - a)y^4z^2 + 2(Ad + Ec)y^3z^3 \\ &+ (E^2 + 2cd - b^2 + 2ab)y^2z^4 + 2Edyz^5 + (d^2 - ab^2)z^6 = 0. \end{aligned} \tag{79}$$

This surface has the tangent t, $y=z=0$, for quintuple line, and intersects the tangent plane, $z=0$, also in the line whose equation is

$$2Ax + (A^2-1)y = 0.$$

The locus of the line from the origin to the Halphen point is found from (66), (46) to be an algebraic cone of the ninth order, and *the locus of the line from the origin to the point of intersection of the osculating conic and osculating cubic* is found from (72), (46), (70) to be also an algebraic cone of the ninth order. The equations of these cones can be written without difficulty, but will not be included here.

The loci of three cubic curves which we have considered will now be discussed. The equation of *the locus of the eight-point nodal cubic* is found, by eliminating n from equations (59), (46), to be

$$\begin{aligned} (z - x^2 - y^2 - xz - az^2)[x(y^2 - bz^2) + Ay^3 + cy^2z + Eyz^2 + dz^3] \\ + z(y^2 - bz^2)^2 = 0. \end{aligned} \tag{80}$$

Therefore this locus is an algebraic surface of the fifth order, which intersects the tangent plane, $z=0$, in the asymptotic tangents, $x^2+y^2=0$, and in the line (74), besides having the tangent t, $y=z=0$, for double line. Moreover, this surface intersects the quadric of Moutard (52) in the asymptotic tangents, and is tangent to the quadric of Moutard along the conics cut on this quadric by the planes (56). *The locus of the eight-point cubic of Sannia* is found from (68), (46) to be an algebraic surface of the eighth order, and *the locus of the osculating cubic* is found from (57), (46), (70) to be an algebraic surface of the twelfth order, but the equations of these surfaces will not be included here.

* Kubota and Su, ibid., p. 300.

There are four points whose loci will now be briefly examined, namely, *the intersection of the flex-ray and the projective normal, the Halphen point, the intersection of the flex-ray and the line from the origin to the Halphen point,* and finally *the intersection of the osculating conic and the osculating cubic.* The parametric equations of the locus of the first point are obtained by setting x and y respectively equal to the coordinates given in (62), by adjoining the equation $z=y/n$, and by then replacing the polynomials V, S, R by their expressions in terms of n given in (77). Thus one finds that the locus of the intersection of the flex-ray and the projective normal is a unicursal curve of order six, as Kubota and Su* have remarked. Similarly, one finds that the loci of the other three points in the order just named are unicursal curves of orders seventeen, twelve, and eighteen. A more detailed study of the properties and relations of the loci mentioned in this section is reserved for another occasion.

4. **Applications to conjugate nets.** The theory developed in the preceding section has interesting connections with the theory of conjugate nets. The coordinates x of a point on a surface referred to a conjugate net in ordinary space, and the coordinates y of the point which is the harmonic conjugate of the point x with respect to the foci of the axis of the point x, satisfy a system of equations† of the form

$$
\begin{aligned}
x_{uu} &= px + \alpha x_u + Ly,\\
x_{uv} &= cx + ax_u + bx_v,\\
x_{vv} &= qx + \delta x_v + Ny.
\end{aligned}
\tag{81}
$$

The ray-points of the net at the point x are given by the formulas

$$
x_1 = x_v - ax, \qquad x_{-1} = x_u - bx. \tag{82}
$$

Some of the invariants of the net are given by

$$
\begin{aligned}
8\mathfrak{B}' &= 4a - 2\delta + (\log r)_v, &\quad r &= N/L,\\
8\mathfrak{C}' &= 4b - 2\alpha - (\log r)_u,\\
H &= c + ab - a_u, &\quad \mathfrak{H} &= c + ab + b_v - \delta_u,\\
K &= c + ab - b_v, &\quad \mathfrak{K} &= c + ab + a_u - \alpha_v.
\end{aligned}
\tag{83}
$$

If the four points x, x_{-1}, x_1, y are used as the vertices of the tetrahedron of reference of a local coordinate system, it is not difficult to calculate by familiar methods a power series expansion for one non-homogeneous co-

* Kubota and Su, ibid., p. 300.

† Lane, *Projective Differential Geometry of Curves and Surfaces*, University of Chicago Press, 1932, p. 138.

ordinate z of a point on the surface in terms of the other two coordinates x, y. Thus one obtains, to terms of the fourth order,

$$z = \frac{1}{2}(Lx^2 + Ny^2) + \frac{4}{3}(L\mathfrak{C}'x^3 + N\mathfrak{B}'y^3) + c_0x^4 + 4c_1x^3y + 4c_3xy^3 + c_4y^4 + \cdots, \tag{84}$$

where the coefficients c_0, c_1, c_3, c_4 are defined by

$$\begin{aligned} c_0 &= \tfrac{1}{3}L\mathfrak{C}'[12\mathfrak{C}' + (\log \mathfrak{C}'r^{1/2})_u], \qquad & 4c_1 &= \tfrac{1}{6}L(H - \mathfrak{H}), \\ c_4 &= \tfrac{1}{3}N\mathfrak{B}'[12\mathfrak{B}' + (\log \mathfrak{B}'r^{-1/2})_v], \qquad & 4c_3 &= \tfrac{1}{6}N(K - \mathfrak{K}), \end{aligned} \tag{85}$$

and the coefficient of x^2y^2 is zero. G. M. Green calculated* an expansion essentially the same as this one to terms of the third order, and showed that the coefficient of x^2y^2 is zero, but apparently did not concern himself with the other terms of the fourth order.

Comparison of the expansion (84) with the expansion (29) leads to interesting results. Equations (13) show that the equations of the planes which must be taken as two faces of a new tetrahedron of reference to reduce (84) to precisely the form of (29) are

$$8L\mathfrak{B}'x - (K - \mathfrak{K})z = 0, \qquad 8N\mathfrak{C}'y - (H - \mathfrak{H})z = 0. \tag{86}$$

The points which are used in place of the ray-points as vertices of the new tetrahedron are the points

$$\frac{K - \mathfrak{K}}{8\mathfrak{B}'}x + x_{-1}, \qquad \frac{H - \mathfrak{H}}{8\mathfrak{C}'}x + x_1, \tag{87}$$

and the line which is used in place of the axis as an edge of the new tetrahedron joins the point x to the point

$$\frac{K - \mathfrak{K}}{8L\mathfrak{B}'}x_{-1} + \frac{H - \mathfrak{H}}{8L\mathfrak{C}'}x_1 + y. \tag{88}$$

It is known† that the ray curves and the axis curves of a non-harmonic conjugate net coincide in case $H = \mathfrak{H}$, $K = \mathfrak{K}$. In this case the points (87) are the ray-points (82), and the point (88) is the point y on the axis. Then the expansions (84), (29) automatically agree to terms of the fourth order.

* Green, *Projective differential geometry of one-parameter families of space curves, and conjugate nets on a curved surface* (Second Memoir), American Journal of Mathematics, vol. 38 (1916), p. 287.

† Lane, ibid., p. 147.

The inverse of the transformation (30) of unit point, with l, m, n given by (32), can be used to change the equations in the notation of §3, based on equation (45), to the notation based on equation (29) with arbitrary unit point. Thus, for example, equation (74) becomes

$$a_2b_0x + a_0b_3y = 0. \tag{89}$$

In the notation of the theory of conjugate nets this equation becomes

$$\mathfrak{C}'x + \mathfrak{B}'y = 0. \tag{90}$$

This is the line joining the point x to the point of intersection of the ray and the associate ray, which has been called* by Davis *the second canonical tangent* of the conjugate net and its associate conjugate net. Thus the following theorem is proved.

The cone of Kubota and Su associated with a tangent of a conjugate net at a point of a surface intersects the tangent plane, besides in the tangent itself, also in Davis's second canonical tangent at the point.

* W. M. Davis, *Contributions to the Theory of Conjugate Nets*, Chicago Doctor's Thesis (1932), p. 19.

University of Chicago,
Chicago, Ill.

CONVERGENCE PROPERTIES OF FOURIER SERIES*

BY

OTTO SZÁSZ

I. Introduction

1. In our previous papers† we were concerned with Fourier series

$$f(\theta) \sim \frac{a_0}{2} + \sum_{\nu=1}^{\infty} (a_\nu \cos \nu\theta + b_\nu \sin \nu\theta)$$

the coefficients of which satisfy the conditions

$$\nu a_\nu \geqq -K, \ \nu b_\nu \geqq -K, \ \nu \geqq 1, \ K \text{ a non-negative constant.}$$

These conditions characterize a special case of "slowly oscillating" series, and it is natural to generalize our results in this direction. In what follows the knowledge of our previous papers is not presupposed, except for the proof of Lemma 5 below.

2. We start with some preliminary notions. A series of real terms $\sum_0^\infty c_\nu$, or the corresponding sequence of its partial sums $\{s_\nu\}$, is called slowly oscillating from below if to any positive ϵ there corresponds an integer $N = N(\epsilon)$ and a positive number $\delta = \delta(\epsilon)$ such that

$$s_{n+k} - s_n > -\epsilon, \ n > N, \ 0 < k < \delta(n+1).$$

It is readily seen that this condition is equivalent to

$$\lim_{\delta \to 0} \liminf_{n \to \infty} \min_{0<k<\delta(n+1)} (s_{n+k} - s_n) \geqq 0.$$

We denote by (A) the class of such series, or sequences.

A series is called slowly oscillating from below in the generalized sense if there exist two positive numbers q and δ such that

$$s_{n+k} - s_n > -q; \quad n = 0, 1, 2, \cdots; \quad 0 < k < \delta(n+1),$$

which is equivalent to

$$\liminf_{n \to \infty} \min_{0<k<\delta(n+1)} (s_{n+k} - s_n) > -\infty.$$

The class of such series will be denoted by $(\overline{A})$. It is obvious that $(A) \subset (\overline{A})$.

---

* Presented to the Society, September 6, 1934; received by the editors August 4, 1934.

† See [5, 6]. The numbers in brackets refer to the list of the author's previous papers at the end of the present paper.

These notions are important for converse theorems of the theory of summability.*

Let (A') be the class of series satisfying the condition $\liminf_{n\to\infty} nc_n > -\infty$. Then we have $(A') \subset (A) \subset (\overline{A})$. If $\{\delta_\nu\}$ is any bounded sequence of non-negative numbers and $\sum_0^\infty c_\nu \subset (A')$ then also $\sum_0^\infty c_\nu\delta_\nu \subset (A')$. This property was used in our papers [5, 6]. An analogous but larger class of series is obtained if we observe that, in order that the series $\sum_0^\infty c_\nu\delta_\nu \subset (A)$ for an arbitrary bounded and non-negative sequence $\{\delta_\nu\}$, it is necessary and sufficient that it belong to (A) in the special case

$$\delta_n = 1, \text{ or } 0, \text{ according as } c_n < 0, \text{ or } \geqq 0,$$

in other words that the series $\sum_0^\infty \frac{1}{2}(c_\nu - |c_\nu|) \subset (A)$. The class of such series we denote by (B). Thus (B) is the class of series such that $\sum_0^\infty c_\nu\delta_\nu \subset (A)$ whenever $\sum_0^\infty c_\nu \subset (A)$. We have the inclusion relation $(A') \subset (B) \subset (A) \subset (\overline{A})$.

Similarly we say that $\sum_0^\infty c_\nu \subset (\overline{B})$ if $\sum_0^\infty \frac{1}{2}(c_\nu - |c_\nu|) \subset (\overline{A})$. The class $(\overline{B})$ is identical to the class of series $\sum_0^\infty c_\nu$ such that $\sum_0^\infty c_\nu\delta_\nu \subset (\overline{A})$ for an arbitrary bounded and non-negative sequence $\{\delta_\nu\}$, whenever $\sum_0^\infty c_\nu \subset (\overline{A})$. It is plain that $(A') \subset (B) \subset (\overline{B}) \subset (\overline{A})$.

3. In the present paper we propose to prove the following theorems.

THEOREM 1. *Let* $\sum_1^\infty a_\nu r^\nu \cos \nu\theta$, $0 \leqq r < 1$, *represent a harmonic function and let*

$$\text{(i)} \qquad \sum_{\nu=1}^{\infty} a_\nu \subset (\overline{B}),$$

$$\text{(ii)} \qquad \sum_{\nu=1}^{\infty} a_\nu r^\nu = O(1), \quad 0 \leqq r < 1.$$

Then also

$$\text{(ii}'\text{)} \qquad \sum_{\nu=1}^{n} a_\nu = O(1) \qquad (n = 1, 2, 3, \cdots).$$

Conversely, the assumption (ii′) *implies* (ii). *If, in addition to* (i) *and* (ii) *we assume that for a fixed* $\theta = \theta_0$

$$\text{(iii)} \qquad \sum_{\nu=1}^{\infty} a_\nu r^\nu \cos \nu\theta_0 = O(1), \quad 0 \leqq r < 1,$$

then also

$$\text{(iii}'\text{)} \qquad \sum_{\nu=1}^{n} a_\nu \cos \nu\theta_0 = O(1) \qquad (n = 1, 2, 3, \cdots).$$

---

* Cf. [2], p. 332; [3], p. 30; [4], p. 326.

Conversely, (iii′) *implies* (iii). *Finally, if*

$$\phi(\theta) \sim \sum_{\nu=1}^{\infty} a_\nu \cos \nu\theta$$

is a Fourier series and

$$\sum_{\nu=1}^{n} a_\nu \cos \nu\theta_0 = O(1) \qquad (n = 1, 2, 3, \cdots),$$

then

$$\int_0^h \{\phi(\theta_0 + t) + \phi(\theta_0 - t)\} dt = O(h) \text{ as } h \to 0.$$

THEOREM 2. *Let* $\sum_1^\infty a_\nu \subset (B)$ *and* $\sum_1^\infty a_\nu r^\nu \to s$ *as* $r \to 1-0$. *Then* $\sum_1^\infty a_\nu$ *converges to* s.* *If in addition, for a fixed* $\theta = \theta_0$,

$$\sum_{\nu=1}^{\infty} a_\nu r^\nu \cos \nu\theta_0 \to s(\theta_0) \text{ as } r \to 1 - 0,$$

then $\sum_1^\infty a_\nu \cos \nu\theta_0$ *converges to* $s(\theta_0)$. *Finally, if*

$$\phi(\theta) \sim \sum_{\nu=1}^{\infty} a_\nu \cos \nu\theta$$

is a Fourier series and $\sum_1^\infty a_\nu \cos \nu\theta_0$ *converges to* $s(\theta_0)$, *then*

$$(2h)^{-1} \int_0^h \{\phi(\theta_0 + t) + \phi(\theta_0 - t)\} dt \to s(\theta_0) \text{ as } h \to 0.$$

THEOREM 3. *Let*

$$\omega(\theta) \sim \sum_{\nu=1}^{\infty} b_\nu \sin \nu\theta$$

be a Fourier series and $\sum_1^\infty b_\nu \subset (\overline{B})$. *Let*

$$\int_0^h \omega(t) dt = O(h) \text{ as } h \to 0.$$

Then

$$\sum_{\nu=1}^{n} \nu b_\nu = O(n) \qquad (n = 1, 2, 3, \cdots).$$

If, in addition,

$$\sum_{\nu=1}^{\infty} b_\nu r^\nu \sin \nu\theta_0 = O(1), \quad 0 \leqq r < 1,$$

* This is a well known result of R. Schmidt.

then

$$\sum_{\nu=1}^{n} b_\nu \sin \nu\theta_0 = O(1) \qquad (n = 1, 2, 3, \cdots).$$

Conversely, if

$$\sum_{\nu=1}^{n} b_\nu \sin \nu\theta_0 = O(1) \qquad (n = 1, 2, 3, \cdots),$$

then

$$\int_0^h \{\omega(\theta_0 + t) + \omega(\theta_0 - t)\} dt = O(h) \text{ as } h \to 0.$$

THEOREM 4. *Let* $\sum_1^\infty b_\nu \subset (B)$ *and*

$$\frac{2}{h}\int_0^h \omega(t)dt \to d \text{ as } h \to 0.^*$$

Then

$$\frac{1}{n}\sum_{\nu=1}^{n} \nu b_\nu \to \frac{d}{\pi} \text{ as } n \to \infty.$$

If, in addition,

$$\sum_{\nu=1}^{\infty} b_\nu r^\nu \sin \nu\theta_0 \to s(\theta_0) \text{ as } r \to 1 - 0,$$

then $\sum_1^\infty b_\nu \sin \nu\theta_0$ *converges to* $s(\theta_0)$. *Conversely, if* $\sum_1^\infty b_\nu \sin \nu\theta_0$ *converges, then*

$$\frac{2}{h}\int_0^h \{\omega(\theta_0 + t) + \omega(\theta_0 - t)\} dt \to s(\theta_0) \text{ as } h \to 0.$$

In proving these theorems we are using methods analogous to those of our papers [5, 6]. Similar theorems hold with respect to the uniform boundedness or uniform convergence in a given interval $\alpha \leqq \theta \leqq \beta$. In the hypotheses of our theorems the sums $\sum_1^\infty \frac{1}{2}(a_\nu - |a_\nu|)$, $\sum_1^\infty \frac{1}{2}(b_\nu - |b_\nu|)$ may be replaced by the sums $\sum_1^\infty \frac{1}{2}(-a_\nu - |a_\nu|)$, $\sum_1^\infty \frac{1}{2}(-b_\nu - |b_\nu|)$ respectively.

Analogous theorems may also be stated for double Fourier series, as well as for almost periodic functions or for Fourier integrals.

The treatment of the cosine series is conspicuously simpler than that of the sine series, the reason being that to the value $\theta = 0$ in the first case there

* The quantity d may be interpreted as the generalized jump of $\omega(\theta)$ at $\theta = 0$.

corresponds the series $\sum_1^\infty a_\nu$ while, in the second case, all terms of the series vanish.*

II. The cosine series

1. **Proof of Theorem 1.** We begin by establishing some lemmas.

Lemma 1. *Put*

$$\sum_{\nu=1}^{n} c_\nu = s_n, \quad \sum_{\nu=1}^{n} \nu c_\nu = v_n, \quad v_0 = s_0 = c_0 = 0,$$

and assume that there exist two numbers, $p>0$ and $\mu>0$, such that

(1) $$s_{n+k} - s_n \geqq -p, \quad 1 \leqq k < 1 + \mu(n+1) \qquad (n = 0, 1, 2, \cdots).$$

Then

(2) $$v_{n+k} - v_n \geqq -p(n+k)$$

and

(3) $$v_n > -p\frac{1+\mu}{\mu}n, \quad n > 0.\dagger$$

Since

$$v_n = \sum_{\nu=0}^{n}(s_n - s_\nu), \quad v_{n+k} = \sum_{\nu=0}^{n+k}(s_{n+k} - s_\nu)$$

we have

$$v_{n+k} - v_n = (n+1)(s_{n+k} - s_n) + \sum_{\nu=n+1}^{n+k}(s_{n+k} - s_\nu) \geqq -p(n+1) - p(k-1),$$

which is the desired inequality (2). Now put

$$n_\nu = [n(1+\mu)^{-\nu}] \qquad (\nu = 0, 1, 2, \cdots),$$

so that

$$v_n = \sum_{\nu=0}^{\infty}(v_{n_\nu} - v_{n_{\nu+1}}),$$

where only a finite number of terms are different from zero. In view of the obvious inequality $n_\nu - n_{\nu+1} < 1 + \mu(1 + n_{\nu+1})$ we have

$$v_n \geqq -p\sum_{\nu=0}^{\infty} n_\nu \geqq -pn\sum_{\nu=0}^{\infty}(1+\mu)^{-\nu} = -p\frac{1+\mu}{\mu}n,$$

* The results of this paper were communicated in the author's seminar during the spring term at the Massachusetts Institute of Technology, and also at the colloquium at Brown University May 18, 1934.

† Cf. [2], p. 333; [4], pp. 326–327.

which proves (3).

LEMMA 2. *Under the assumptions* (1) *and*

$$\left|\sum_{\nu=1}^{\infty} c_\nu x^\nu\right| \leqq M, \quad 0 \leqq x < 1, \tag{4}$$

we have

$$|s_n| < M(1+8e)\frac{2+\mu}{\mu} + p\left(1 + 4e\frac{(2+\mu)(1+\mu)}{\mu^2}\right).$$

We set

$$\sum_{\nu=1}^{\infty} c_\nu x^\nu = P(x);$$

then from

$$\frac{v_n}{n+1} = s_n - (n+1)^{-1}\sum_{\nu=0}^{n} s_\nu$$

it follows that

$$\sum_{\nu=1}^{\infty} \frac{v_\nu}{\nu+1} x^\nu = (1-x)^{-1}P(x) - x^{-1}\int_0^x P(t)(1-t)^{-2}dt,$$

and, by (4),

$$\left|\sum_{\nu=1}^{\infty} \frac{v_\nu}{\nu+1} x^\nu\right| \leqq 2M(1-x)^{-1}.$$

Hence

$$\sum_{\nu=1}^{\infty}\left(\frac{v_\nu}{\nu+1} + p\frac{1+\mu}{\mu}\right)x^\nu \leqq 2M(1-x)^{-1} + p\frac{1+\mu}{\mu}x(1-x)^{-1}.$$

The coefficients of the power series of the left-hand member are positive, by (3). Consequently, for $x = 1 - 1/(n+1)$,

$$\begin{aligned}\left(1-\frac{1}{n+1}\right)^n \sum_{\nu=1}^{n}\left(\frac{v_\nu}{\nu+1} + p\frac{1+\mu}{\mu}\right) &< \sum_{\nu=1}^{n}\left(\frac{v_\nu}{\nu+1} + p\frac{1+\mu}{\mu}\right)\left(1-\frac{1}{n+1}\right)^\nu \\ &< \left(2M + p\frac{1+\mu}{\mu}\right)(n+1),\end{aligned}$$

whence

$$\sum_{\nu=1}^{n} \frac{v_\nu}{\nu+1} < \left(2M + p\frac{1+\mu}{\mu}\right)\left(1+\frac{1}{n}\right)^n (n+1)$$
$$< \left(2M + p\frac{1+\mu}{\mu}\right) e(n+1).$$

On the other hand, again by (3),

$$\sum_{\nu=1}^{n} \frac{v_\nu}{\nu+1} > -pn\frac{1+\mu}{\mu}, \qquad n > 0,$$

which yields the result

$$\left|\sum_{\nu=1}^{n} \frac{v_\nu}{\nu+1}\right| < e\left(2M + p\frac{1+\mu}{\mu}\right)(n+1), \qquad n > 0.$$

Now, on setting

$$(n+1)^{-1}\sum_{\nu=0}^{n} s_\nu = \sigma_n, \qquad \sum_{\nu=1}^{n} \frac{v_\nu}{\nu+1} = u_n,$$

we have

$$x^{-1}\int_0^x P(t)(1-t)^{-2}dt = \sum_{\nu=1}^{\infty} \sigma_\nu x^\nu,$$

$$(1-x)x^{-1}\int_0^x P(t)(1-t)^{-2}dt = \sum_{\nu=1}^{\infty} (\sigma_\nu - \sigma_{\nu-1})x^\nu = \sum_{\nu=1}^{\infty} \frac{v_\nu}{\nu(\nu+1)} x^\nu,$$

and*

$$(1-x)x^{-1}\int_0^x P(t)(1-t)^{-2}dt - \sigma_n = (1-x)\sum_{\nu=1}^{\infty} \frac{u_\nu}{\nu+1} x^\nu - \frac{u_n}{n+1}$$
$$+ \sum_{\nu=1}^{n} \frac{u_\nu}{\nu(\nu+1)}(x^\nu - 1) + \sum_{\nu=n+1}^{\infty} \frac{u_\nu}{\nu(\nu+1)} x^\nu.$$

By (4)

$$\left| (1-x)x^{-1}\int_0^x P(t)(1-t)^{-2}dt \right| \leqq M.$$

On combining this with the preceding inequality we have

$$|\sigma_n| < M + 2e\left(2M + p\frac{1+\mu}{\mu}\right)$$
$$+ e\left(2M + p\frac{1+\mu}{\mu}\right)\left\{n(1-x) + \frac{1}{n+1}\frac{x^{n+1}}{1-x}\right\},$$

* Cf. [1], Theorem 1, with νa_ν replaced by $v_\nu/(\nu+1)$.

whence, for $x=1-1/(n+1)$,

$$|\sigma_n| < M + 4e\left(2M + p\frac{1+\mu}{\mu}\right) \equiv M_1, \qquad n \geqq 1.$$

Consequently

$$s_n = \frac{v_n}{n+1} + \sigma_n > -p\frac{1+\mu}{\mu} - M_1, \qquad n \geqq 1.$$

On the other hand we have*

$$s_n = \sigma_{n+k} + \frac{n+1}{k}(\sigma_{n+k} - \sigma_n) - \frac{1}{k}\sum_{\nu=1}^{k}(s_{n+\nu} - s_n),$$

whence, for $\mu(n+1) \leqq k < 1+\mu(n+1)$,

$$s_n < M_1 + 2M_1/\mu + p.$$

Lemma 2 follows at once by combining these two inequalities for s_n.

We now pass on to the proof of Theorem 1. Under the hypotheses (i) and (ii) Lemma 2 implies

$$s_n = \sum_{\nu=1}^{n} a_\nu = O(1) \qquad (n = 1, 2, 3, \cdots).$$

Conversely, if this condition is satisfied then

$$\sum_{\nu=1}^{\infty} a_\nu x^\nu = (1-x)\sum_{\nu=1}^{\infty} s_\nu x^\nu = O(1), \qquad 0 \leqq x < 1.$$

Furthermore, the series $\sum_1^\infty a_\nu(1-\cos\nu\theta_0)$ also $\subset (\overline{A})$. Then, if (iii) is satisfied, by Lemma 2,

$$\sum_{\nu=1}^{n} a_\nu(1-\cos\nu\theta_0) = O(1), \text{ and } \sum_{\nu=1}^{n} a_\nu \cos\nu\theta_0 = O(1) \qquad (n = 1, 2, 3, \cdots).$$

The converse is proved by the same argument as before.

To prove the last statement of the theorem we observe that

$$(2h)^{-1}\int_0^h \{\phi(\theta_0 + t) + \phi(\theta_0 - t)\}dt = \sum_{\nu=1}^{\infty} a_\nu \cos\nu\theta_0(\sin\nu h)/(\nu h).$$

By assumption $\sum_1^\infty(a_\nu - |a_\nu|) \subset (\overline{A})$, and we have proved that $\sum_1^n a_\nu = O(1)$. Hence the series $-\sum_1^\infty a_\nu$ as well as $-\sum_1^\infty |a_\nu| \subset (\overline{A})$. This means that, for suitably chosen p and μ,

$$\sum_{\nu=n+1}^{n+k} |a_\nu| \leqq p, \qquad 1 \leqq k < 1 + \mu(n+1), \qquad n \geqq 0.$$

* Cf. [3], p. 31.

Then, by Lemma 1,

$$\sum_{\nu=1}^{n} \nu \mid a_\nu \mid < np \frac{1+\mu}{\mu}, \qquad n > 0.$$

Now put

$$\sum_{\nu=1}^{\infty} a_\nu \cos \nu\theta_0 (\sin \nu h)/(\nu h) = \sum_{\nu=1}^{n} a_\nu \cos \nu\theta_0 + \sum_{\nu=1}^{n} a_\nu \cos \nu\theta_0((\sin \nu h)/(\nu h) - 1)$$
$$+ \sum_{\nu=n+1}^{\infty} a_\nu \cos \nu\theta_0 (\sin \nu h)/(\nu h) \equiv S_1 + S_2 + S_3.$$

In the previous argument we have proved the existence of a constant G such that

$$\mid S_1 \mid \leqq G.$$

From

$$0 \leqq 1 - (\sin \nu h)/(\nu h) \leqq \tfrac{1}{6}\nu^2 h^2$$

we have

$$\mid S_2 \mid \leqq \tfrac{1}{6} h^2 \sum_{\nu=1}^{n} \nu^2 \mid a_\nu \mid \leqq \tfrac{1}{6} n h^2 \sum_{\nu=1}^{n} \nu \mid a_\nu \mid < p \frac{1+\mu}{6\mu} n^2 h^2. \tag{5}$$

Finally, on writing $\tau_{nk} = \sum_{n+1}^{n+k} \mid a_\nu \mid / \nu$ we have

$$\tau_{nk} \leqq (n+1)^{-1} \sum_{\nu=n+1}^{n+k} \mid a_\nu \mid \leqq p/(n+1), \qquad 1 \leqq k < 1 + \mu(n+1).$$

Let $n_0, n_1, \cdots$ be a sequence of integers such that

$$(1+\mu)n_{\nu-1} + \mu \leqq n_\nu < (1+\mu)(1+n_{\nu-1}), \qquad n_0 = n.$$

Then

$$\sum_{\nu=n+1}^{n_k} \mid a_\nu \mid / \nu = \sum_{\nu=n_0+1}^{n_1} + \cdots + \sum_{\nu=n_{k-1}+1}^{n_k} \leqq p \sum_{\nu=0}^{k-1} (n_\nu + 1)^{-1}$$
$$\leqq \frac{p}{n+1} \sum_{\nu=0}^{k-1} (1+\mu)^{-\nu} < \frac{p(1+\mu)}{(n+1)\mu},$$

since

$$n_\nu + 1 \geqq (1+\mu)(n_{\nu-1}+1) \geqq (1+\mu)^\nu (n+1).$$

Thus

$$\sum_{\nu=n+1}^{\infty} \mid a_\nu \mid / \nu \leqq \frac{p(1+\mu)}{(n+1)\mu}$$

and

$$(6) \qquad |S_3| \leqq h^{-1} \sum_{\nu=n+1}^{\infty} |a_\nu|/\nu \leqq \frac{p(1+\mu)}{h(n+1)\mu}.$$

On choosing $n=[1/h]$, $h<1$, we finally obtain

$$\left| (2h)^{-1} \int_0^h \{\phi(\theta_0+t)+\phi(\theta_0-t)\} dt \right| < G + \frac{7}{6} p \frac{1+\mu}{\mu}.$$

2. **Proof of Theorem 2.** By hypothesis $\sum_1^\infty (a_\nu - |a_\nu|) \subset (A)$. Since $\sum_1^\infty a_\nu$ converges, $-\sum_1^\infty |a_\nu|$ also $\subset (A)$. Hence $\sum_1^\infty a_\nu \cos \nu\theta \subset (A)$ for every θ. If now

$$\sum_{\nu=1}^{\infty} a_\nu r^\nu \cos \nu\theta_0 \to s(\theta_0) \quad \text{as} \quad r \to 1-0,$$

the theorem of R. Schmidt mentioned above implies

$$\sum_{\nu=1}^{\infty} a_\nu \cos \nu\theta_0 = s(\theta_0).$$

It remains to prove the last statement of Theorem 2. We write

$$\begin{aligned}
(2h^{-1}) \int_0^h & \{\phi(\theta_0+t)+\phi(\theta_0-t)\} dt - s(\theta_0) \\
&= \left(\sum_{\nu=1}^{n} a_\nu \cos \nu\theta_0 - s(\theta_0) \right) + \sum_{\nu=1}^{n} a_\nu \cos \nu\theta_0 \left(\frac{\sin \nu h}{\nu h} - 1 \right) \\
&\quad + \sum_{\nu=n+1}^{\lambda_n} a_\nu \cos \nu\theta_0 \frac{\sin \nu h}{\nu h} + \sum_{\nu=\lambda_n+1}^{\infty} a_\nu \cos \nu\theta_0 \frac{\sin \nu h}{\nu h} \\
&= Z_0 + Z_1 + Z_2 + Z_3,
\end{aligned}$$

where $\lambda_n > n$ will be fixed later. For a given ϵ, $0<\epsilon<1$, choose $N=N(\epsilon)$ so that

$$(7) \qquad |Z_0| < \epsilon^2, \qquad n > N.$$

From (5) and (6) we have

$$|Z_1| < p \frac{1+\mu}{\mu} n^2 h^2, \qquad |Z_3| \leqq \frac{p(1+\mu)}{h(\lambda_n+1)\mu}.$$

Let $h_0 = h_0(\epsilon)$ be so small that

$$h_0 < \epsilon, \qquad [\epsilon/h_0] > N,$$

and put

$$n = [\epsilon/h], \qquad \lambda_n = [\pi/(\epsilon h)], \qquad h < h_0,$$

so that

$$nh \leqq \epsilon, \qquad 1 + \lambda_n > \pi/(\epsilon h).$$

Then

$$|Z_1| < p\,\frac{1+\mu}{\mu}\,\epsilon^2, \qquad |Z_3| < \frac{p(1+\mu)}{\pi\mu}\,\epsilon.$$

To estimate Z_2 subdivide the range $0<\nu h\leqq\lambda_n h<\pi/\epsilon$ into subintervals in each of which $(\sin \nu h)/(\nu h)$ is monotone (as function of ν), the number of these subintervals being not greater than $1+[1/\epsilon]$. An easy application of the partial summation formula together with (7) will show that

$$|Z_2| < (2/\epsilon)2\epsilon^2 = 4\epsilon.$$

To complete the proof of Theorem 2 it remains to allow $\epsilon\to 0$.

III. The sine series

1. **Proof of Theorem 3.** We shall need some additional lemmas.

LEMMA 3. *Let*

$$\omega(\theta) \sim \sum_{\nu=1}^{\infty} b_\nu \sin \nu\theta.$$

If

$$\int_0^h \omega(t)dt = O(h) \ \textit{as}\ h \to 0, \tag{8}$$

then

$$\sum_{\nu=1}^{\infty} \nu b_\nu r^\nu = O((1-r)^{-1}) \ \textit{as}\ r \to 1-0.$$

We have*

$$\sum_{\nu=1}^{\infty} \nu b_\nu r^\nu = -\pi^{-1}\int_0^\pi \omega(t)\,\frac{d}{dt}\,p(r,t)dt, \qquad 0 \leqq r < 1,$$

where

$$p(r,t) = (1-r^2)/\Delta, \quad \Delta = 1 - 2r\cos t + r^2 = (1-r)^2 + 4r\sin^2\frac{t}{2},$$

$$\frac{d}{dt}\,p(r,t) = -2r(1-r^2)\sin t/\Delta^2.$$

On setting

$$\psi(t) = \int_0^t \omega(\tau)d\tau$$

* Cf. [5], formula (21).

and integrating by parts we have

$$\sum_{\nu=1}^{\infty} \nu b_\nu r^\nu = -\pi^{-1}2r(1-r^2)\int_0^\pi \psi(t)\frac{d}{dt}(\sin t\,\Delta^{-2})dt.$$

In view of (8) an easy calculation yields

$$\sum_{\nu=1}^{\infty} \nu b_\nu r^\nu = O\left\{(1-r)^2\int_0^\infty t\{(1-r)^2+t^2\}^{-2}dt\right\} = O((1-r)^{-1}).$$

LEMMA 4. *If* (8) *holds and* $\sum_1^\infty b_\nu \subset (\overline{A})$, *then*

$$v_n = \sum_{\nu=1}^{n} \nu b_\nu = O(n) \qquad (n = 1, 2, 3, \cdots).$$

By Lemma 3

$$\sum_{\nu=1}^{\infty} \nu b_\nu r^\nu = O((1-r)^{-1}),$$

while, by Lemma 1,

$$v_\nu + p\frac{1+\mu}{\mu}\nu > 0 \qquad (\nu = 1, 2, 3, \cdots), \tag{3'}$$

with suitably chosen p and μ. Hence

$$\sum_{\nu=1}^{\infty}\left(v_\nu + p\frac{1+\mu}{\mu}\nu\right)r^\nu = O((1-r)^{-2}),$$

$$\sum_{\nu=1}^{n}\left(v_\nu + p\frac{1+\mu}{\mu}\nu\right) = O\left\{\sum_{\nu=1}^{n}\left(v_\nu + p\frac{1+\mu}{\mu}\nu\right)\left(1-\frac{1}{n}\right)^\nu\right\} = O(n^2).$$

Thus

$$V_n = \sum_{\nu=1}^{n} v_\nu = O(n^2) \qquad (n = 1, 2, 3, \cdots).$$

On the other hand the relation

$$kv_n = \frac{kV_{n+k}}{n+k+1} + (n+1)\left(\frac{V_{n+k}}{n+k+1} - \frac{V_n}{n+1}\right) - \sum_{\nu=1}^{k}(v_{n+\nu} - v_n)$$

gives, in view of (2),

$$v_n < Cn,$$

where C is a generic notation for a constant, not necessarily the same in all formulas where it occurs. On combining this with (3′) we obtain a proof of Lemma 4, and also of the first statement of Theorem 3.

Now, if $\sum_1^\infty b_\nu \subset (\overline{B})$, then $\sum_1^\infty b_\nu(1-\sin \nu\theta) \subset (\overline{A})$. If, in addition,

$$\sum_{\nu=1}^{\infty} b_\nu r^\nu \sin \nu\theta_0 = O(1), \qquad 0 \leqq r < 1,$$

then, as in the proof of Lemma 2,

$$\sum_{\nu=1}^{\infty} (\nu+1)^{-1}(b_1 \sin \theta_0 + \cdots + \nu b_\nu \sin \nu\theta_0) r^\nu = O((1-r)^{-1}),$$

while, by Lemma 4,

$$\sum_{\nu=1}^{\infty} \nu b_\nu r^\nu = O((1-r)^{-1}), \qquad \sum_{\nu=1}^{\infty} (b_1 + \cdots + \nu b_\nu) r^\nu = O((1-r)^{-2}),$$

and

$$\sum_{\nu=1}^{\infty} (\nu+1)^{-1}(b_1 + \cdots + \nu b_\nu) r^\nu = O((1-r)^{-1}).$$

Thus

$$\sum_{\nu=1}^{\infty} (\nu+1)^{-1}\{b_1(1-\sin\theta_0) + \cdots + \nu b_\nu(1-\sin \nu\theta_0)\} r^\nu = O((1-r)^{-1}).$$

We can now apply Lemma 1 obtaining

$$S_n = \sum_{\nu=1}^{n} \nu b_\nu (1 - \sin \nu\theta_0) > -Cn, \qquad n \geqq 1. \tag{9}$$

Using the same argument as in the proof of Lemma 4 we get

$$R_n = \sum_{\nu=1}^{n} S_\nu/(\nu+1) = O(n) \qquad (n = 1, 2, 3, \cdots).$$

Writing for simplicity $R_0 = 0$ we have

$$\begin{aligned}\sum_{\nu=1}^{n} S_\nu &= \sum_{\nu=1}^{n} (\nu+1)S_\nu/(\nu+1) = \sum_{\nu=1}^{n} (\nu+1)(R_\nu - R_{\nu-1}) \\ &= (n+1)R_n - \sum_{\nu=1}^{n-1} R_\nu = O(n^2),\end{aligned}$$

whence, again as in Lemma 4,

$$S_n < Cn.$$

On combining these results we have

$$\sum_{\nu=1}^{n} \nu b_\nu \sin \nu\theta_0 = O(n) \qquad (n = 1, 2, 3, \cdots).$$

This, together with

$$\sum_{\nu=1}^{\infty} b_\nu r^\nu \sin \nu\theta_0 = O(1), \qquad 0 \leqq r < 1,$$

gives

$$\sum_{\nu=1}^{n} b_\nu \sin \nu\theta_0 = O(1) \qquad (n = 1, 2, 3, \cdots),$$

by an argument analogous to that used in the proof of Lemma 2.

Conversely, assume

$$\sum_{\nu=1}^{n} b_\nu \sin \nu\theta_0 = O(1) \qquad (n = 1, 2, 3, \cdots),$$

and write

$$(2h)^{-1}\int_0^h \left\{\omega(\theta_0 + t) + \omega(\theta_0 - t)\right\} dt = \sum_{\nu=1}^{\infty} b_\nu \sin \nu\theta_0 \frac{\sin \nu h}{\nu h}$$

$$= \sum_{\nu=1}^{n} b_\nu \sin \nu\theta_0 + \sum_{\nu=1}^{n} b_\nu \sin \nu\theta_0 \left(\frac{\sin \nu h}{\nu h} - 1\right) + \sum_{\nu=n+1}^{\infty} b_\nu \sin \nu\theta_0 \frac{\sin \nu h}{\nu h}$$

$$\equiv J_1 + J_2 + J_3.$$

Then, by assumption,

$$J_1 = O(1).$$

For J_2 we obtain the estimate

$$|J_2| \leqq \tfrac{1}{6}h^2 \sum_{\nu=1}^{n} \nu^2 |b_\nu| \leqq \tfrac{1}{6}nh^2 \sum_{\nu=1}^{n} \nu |b_\nu|.$$

But, by Lemma 4, $\sum_1^n \nu b_\nu = O(n)$ and, since $\sum_1^\infty b_\nu \subset (\overline{B})$, by Lemma 1, $\sum_1^n \nu(b_\nu - |b_\nu|) > -Cn$. Thus

$$g_n = \sum_{\nu=1}^{n} \nu |b_\nu| = O(n)$$

and

$$J_2 = h^2 O(n^2).$$

Finally, on putting $t_{nk} = \sum_{n+1}^{n+k} |b_\nu|/\nu$ we have

$$t_{nk} = \sum_{\nu=n+1}^{n+k} (g_\nu - g_{\nu-1})\nu^{-2}$$

$$= - g_n(n+1)^{-2} + g_{n+k}(n+k)^{-2} + \sum_{\nu=n+1}^{n+k-1} g_\nu(\nu^{-2} - (\nu+1)^{-2})$$

$$= O\left(\frac{1}{n}\right).$$

Hence $\sum_{n+1}^{\infty} |b_\nu|/\nu = O(1/n)$ and

$$|J_3| \leqq h^{-1} \sum_{\nu=n+1}^{\infty} |b_\nu|/\nu = h^{-1}O\left(\frac{1}{n}\right).$$

On choosing $n=[1/h]$, $h<1$, we get immediately

$$\int_0^h \{\omega(\theta_0+t) + \omega(\theta_0-t)\} dt = O(h).$$

2. **Proof of Theorem 4.** We first state two additional lemmas.

LEMMA 5. *Let*

$$\omega(\theta) \sim \sum_{\nu=1}^{\infty} b_\nu \sin \nu\theta.$$

If the limit

$$\lim_{h\to 0} \frac{2}{h}\int_0^h \omega(t)dt = d \tag{10}$$

exists, then

$$(1-r)\sum_{\nu=1}^{\infty} \nu b_\nu r^\nu \to \frac{d}{\pi} \quad as \quad r \to 1 - 0.*$$

LEMMA 6. *If* (10) *holds and* $\sum_1^\infty b_\nu \subset (A)$, *then*

$$\frac{v_n}{n} = n^{-1}\sum_{\nu=1}^{n} \nu b_\nu \to \frac{d}{\pi} \quad as \quad n \to \infty.$$

By Lemma 5,

$$(1-r)\sum_{\nu=1}^{\infty} \nu b_\nu r^\nu \to \frac{d}{\pi} \text{ as } r \to 1,$$

while, by Lemma 1,

$$v_\nu + p\frac{1+\mu}{\mu}\nu > 0 \qquad (\nu = 1, 2, 3, \cdots).$$

* This was proved in our paper [5], Theorem 5.

with suitably chosen p and μ. Hence

$$(1-r)^2\sum_{\nu=1}^{\infty} v_\nu r^\nu \to \frac{d}{\pi},$$

$$(1-r)^2\sum_{\nu=1}^{\infty}\left(v_\nu + p\frac{1+\mu}{\mu}\nu\right)r^\nu \to \frac{d}{\pi} + p\frac{1+\mu}{\mu}, \text{ as } r\to 1.$$

By a theorem of Hardy and Littlewood,

$$\sum_{\nu=1}^{n}\left(v_\nu + p\frac{1+\mu}{\mu}\nu\right) \sim \left(\frac{d}{\pi} + p\frac{1+\mu}{\mu}\right)\frac{n^2}{2},$$

so that

$$V_n = \sum_{\nu=1}^{n} v_\nu \sim dn^2/(2\pi) \text{ as } n\to\infty.$$

On the other hand the relation

$$v_n/n = (m-n)^{-1}(V_m/m - V_n/n) + V_m/(mn) - n^{-1}(m-n)^{-1}\sum_{\nu=n+1}^{m}(v_\nu - v_n),$$
$$m > n > 0,$$

combined with

$$V_n = n^2(d/(2\pi) + \epsilon_n), \qquad \epsilon_n \to 0,$$

gives

$$\begin{aligned} v_n/n &= (m-n)^{-1}\{(m-n)d/(2\pi) + m\epsilon_m - n\epsilon_n\} + m(d/(2\pi)+\epsilon_m)/n \\ &\quad - \sum_{\nu=n+1}^{m}(v_\nu - v_n)/(n(m-n)) \\ &= (1+m/n)d/(2\pi) + \epsilon_m m^2/(n(m-n)) \\ &\quad - \epsilon_n n/(m-n) - \sum_{\nu=n+1}^{m}(v_\nu - v_n)/(n(m-n)). \end{aligned}$$

Here we may use inequality (2) of Lemma 1 with

$$p = \epsilon, \quad \mu(n+1) + n \leqq m < (\mu+1)(n+1), \quad \mu = \mu(\epsilon)\to 0 \text{ as } \epsilon\to 0,$$

which gives

$$\limsup_{n\to\infty} v_n/n \leqq (2+\mu)d/(2\pi) + (1+\mu)\epsilon.$$

On allowing here $\epsilon\to 0$ and $\mu\to 0$, we get

$$\limsup_{n\to\infty} v_n/n \leqq d/\pi.$$

In a similar fashion the relation

$$v_m/m = (m-n)^{-1}(V_m/m - V_n/n) + V_n/(nm) + m^{-1}(m-n)^{-1}\sum_{\nu=n+1}^{m}(v_m - v_\nu)$$

yields

$$\liminf_{m\to\infty} v_m/m \geqq d/\pi,$$

whence

$$\lim_{n\to\infty} v_n/n = d/\pi.$$

This proves Lemma 6, and also the first statement of Theorem 4. Assume now

$$\sum_{\nu=1}^{\infty} b_\nu r^\nu \sin \nu\theta_0 \to s(\theta_0) \text{ as } r \to 1-0.$$

Then it is readily seen that

$$(1-r)\sum_{\nu=1}^{\infty}(\nu+1)^{-1}(b_1 \sin\theta_0 + \cdots + \nu b_\nu \sin \nu\theta_0)r^\nu \to 0 \text{ as } r\to 1;$$

hence

$$(1-r)\sum_{\nu=1}^{\infty}(\nu+1)^{-1}(b_1(1-\sin\theta_0) + \cdots + \nu b_\nu(1-\sin\nu\theta_0))r^\nu \to d/\pi \text{ as } r\to 1.$$

Being combined with (9) this yields

$$R_n = \sum_{\nu=1}^{n} S_\nu/(\nu+1) \sim dn/\pi.$$

Taking into account that

$$\sum_1^\infty b_\nu(1-\sin \nu\theta_0) \subset (A)$$

and using the same argument as in the proof of Lemma 6 we conclude

$$S_n \to d/\pi, \text{ or } n^{-1}\sum_{\nu=1}^{n}\nu b_\nu \sin\nu\theta_0 \to 0 \text{ as } n\to\infty.$$

The relation

$$\sum_1^\infty b_\nu \sin \nu\theta_0 = s(\theta_0)$$

now follows by the classical theorem of Tauber.

Conversely, assuming that this relation is satisfied, we have

$$
\begin{aligned}
&\frac{2}{h}\int_0^h \{\omega(\theta_0+t)+\omega(\theta_0-t)\}dt - s(\theta_0)\\
&\quad = \left(\sum_{\nu=1}^{n} b_\nu \sin \nu\theta_0 - s(\theta_0)\right) + \sum_{\nu=1}^{n} b_\nu \sin \nu\theta_0 \left(\frac{\sin \nu h}{\nu h} - 1\right)\\
&\qquad + \sum_{\nu=n+1}^{\lambda_n} b_\nu \sin \nu\theta_0 \frac{\sin \nu h}{\nu h} + \sum_{\nu=\lambda_n+1}^{\infty} b_\nu \sin \nu\theta_0 \frac{\sin \nu h}{\nu h}\\
&\quad \equiv U_0 + U_1 + U_2 + U_3.
\end{aligned}
$$

Given ϵ, $0<\epsilon<1$, we first choose $N=N(\epsilon)$ so that

$$| U_0 | < \epsilon^2, \qquad n > N.$$

We also have

$$| U_1 | \leqq \tfrac{1}{6}nh^2 \sum_{\nu=1}^{n} \nu | b_\nu | < Cn^2h^2,$$

$$| U_3 | \leqq h^{-1} \sum_{\nu=\lambda_n+1}^{\infty} | b_\nu | /\nu < C/(h\lambda_n).$$

Now choose $n=[\epsilon/h]$, $\lambda_n=[\pi/(\epsilon h)]$. An estimate for U_2 and the final result

$$\frac{2}{h}\int_0^h \{\omega(\theta_0+t)+\omega(\theta_0-t)\}dt \to s(\theta_0) \text{ as } h\to 0$$

is obtained by precisely the same argument as in the proof of Theorem 2.

References to previous papers by the author

1. *Verallgemeinerung eines Littlewoodschen Satzes über Potenzreihen*, Journal of the London Mathematical Society, vol. 3 (1928), pp. 256–262.

2. *Verallgemeinerung und neuer Beweis einiger Sätze Tauberscher Art*, Sitzungsberichte der Mathematisch-Physikalischen Klasse der Akademie der Wissenschaften zu München, 1929, pp. 325–340.

3. *Über Sätze Tauberscher Art*, Jahresbericht der Deutschen Mathematiker Vereinigung, vol. 39 (1930), pp. 28–31.

4. *Über einige Sätze von Hardy und Littlewood*, Göttinger Nachrichten, 1930, pp. 315–333.

5. *Zur Konvergenztheorie der Fourierschen Reihen*, Acta Mathematica, vol. 61 (1933), pp. 185–201.

6. *A Fourier-féle sor részletösszegeinek korlátosságáról és összetartásáról*, Mathematischer und Naturwissenschaftlicher Anzeiger der Ungarischen Akademie der Wissenschaften, vol. 50 (1933), pp. 125–146.

Massachusetts Institute of Technology,
Cambridge, Mass.